Niveau(x) d'enseignement: 0 Numéro EDUQ : 8570163
Centre d'intérêt: 540

Dossier sur l'évaluation / présenté par France Fontaine. -- Montréal : Université de Montréal, Service pédagogique, 1979.
138 p. ; 28 cm.

Comprend des bibliogr. Dépouillement: Evaluer? Mesurer? / France Fontaine -- Assurer à tous la compétence / Claire Fel.-Larivière, France Fontaine -- Les modèles d'évaluation de l'apprentissage ; Il ne faut pas confondre l'autonotation et l'auto-évaluation / France Fontaine -- L'utilisation des objectifs pour évaluer l'apprentissage lors d'un stage / Huguette Bernard, France Fontaine -- Démarche pour évaluer l'enseignement ; Les tests sur les prérequis et les pré-tests ; Faut-il pénaliser le "choix au hasard?" / France Fontaine -- Les systèmes de notation / Huguette Bernard -- Abolir les notes ; pourquoi pas? / France Fontaine -- Un système informatique de contrôle de l'apprentissage / Gilles Génier -- La banque d'objectifs et d'items : un exemple d'application / Huguette Bernard, Serge Normand -- L'évaluation de la formation clinique en ergothérapie / Francine Ferland, Raymonde Hachey.

ISBN 2-920096-02-8

Descripteur(s): *Evaluation en éducation *Méthode d'évaluation *Auto-évaluation *Evaluation de l'apprentissage *Instrument d'évaluation *Evaluation de tests *Note

---

Le dossier a pour but d'appuyer les professeurs dans leurs réflexions et leurs démarches relatives à l'évaluation. Il réunit divers textes sans pour autant constituer une revue de la documentation ni un tour d'horizon complet sur le sujet. Il s'agit d'un outil de référence auquel le professeur peut facilement se référer, et qui est conçu de façon à pouvoir s'enrichir de textes complémentaires.

Le dossier est divisé en cinq sections. La première regroupe les textes définissant les concepts d'évaluation; on trouve dans la seconde des textes présentant les processus d'élaboration de deux instruments d'évaluation différents; par la suite, on présente des textes se rapportant à des pratiques particulières d'évaluation de l'apprentissage; la quatrième section rassemble les travaux sur la notation; enfin, la dernière section regroupe trois expériences d'évaluation.

Type de document: d

EDUQ 8570163

Dossier sur l'évaluation

présenté par France Fontaine

Service pédagogique

Université de Montréal

EDUQ 8570163

Dossier sur l'évaluation

Présenté par
France Fontaine

Service Pédagogique
Université de Montréal
1979

Auteurs
Huguette Bernard, conseiller pédagogique,
Service pédagogique, Université de Montréal

Claire Felx-Larivière, conseiller pédagogique,
Service pédagogique, Université de Montréal

Francine Ferland, professeur,
Ecole de Réadaptation, Université de Montréal

France Fontaine, conseiller pédagogique,
Service pédagogique, Université de Montréal

Gilles Génier, professeur,
Ecole polytechnique de Montréal

Raymonde Hachey, professeur,
Ecole de Réadaptation, Université de Montréal

Serge Normand, conseiller pédagogique,
Service pédagogique, Université de Montréal

Collaboration spéciale
Claude Schnéegans, Chef de la Section Applications,
Centre de calcul, Université de Montréal

Serge Normand, conseiller pédagogique,
Service pédagogique, Université de Montréal

Maquette de la couverture
Vincent Deguire

Dactylographie
Danielle Leroux, Service pédagogique
Lucie Perreault, Service pédagogique

ISBN 2-920096-02-8
Dépôt légal, 2 trimestre 1979
Bibliothèque nationale du Québec
Tous droits de reproduction,
d'adaptation ou de traduction réservés
Service pédagogique, Université de Montréal, 1979

TABLE DES MATIÈRES

PRÉSENTATION DES TEXTES 3

Section 1

La définition des concepts d'évaluation 3

Évaluer? Mesurer?
France Fontaine 5

Assurer à tous la compétence
Claire Felx-Larivière
France Fontaine 11

Les modèles d'évaluation
de l'apprentissage
France Fontaine 17

Il ne faut pas confondre l'autonotation et l'auto-évaluation
France Fontaine 29

Section 2

Des processus d'élaboration d'instrument d'évaluation 33

L'utilisation des objectifs pour évaluer l'apprentissage
lors d'un stage
Huguette Bernard
France Fontaine 35

Démarche pour évaluer l'enseignement
France Fontaine 51

Section 3

Des pratiques d'évaluation de l'apprentissage 73

Les tests sur les prérequis
et les pré-tests
France Fontaine 75

Faut-il pénaliser le «choix au hasard»?
France Fontaine 85

Section 4

La notation 91

Les systèmes de notation
Huguette Bernard 93

Abolir les notes : pourquoi pas?
France Fontaine 109

Section 5

Des expériences en évaluation 115

Un système informatique de contrôle
de l'apprentissage
Gilles Génier 117

La banque d'objectifs et d'items :
un exemple d'application
Huguette Bernard
Serge Normand 123

L'évaluation de la formation clinique en ergothérapie
Francine Ferland
Raymonde Hachey 131

PRÉSENTATION DES TEXTES

L'évaluation est une porte d'entrée privilégiée pour amorcer une réflexion sur l'enseignement. Même si, dans l'élaboration ou la transformation d'un cours ou d'un programme de formation, l'évaluation n'est pas le premier point à traiter, l'évaluation offre souvent l'occasion d'une prise de conscience par les questions qu'elle soulève. L'étudiant n'est-il pas en droit de savoir ce sur quoi on l'évalue et comment on l'évalue? C'est parfois à partir de ce point très «sensible» que s'amorcent une remise en question puis des changements dans l'enseignement.

La publication d'un dossier sur l'évaluation a donc sa place au Service pédagogique. Il a pour but d'appuyer les professeurs dans leurs réflexions et leurs démarches relatives à l'évaluation.

Le dossier réunit divers textes concernant ce thème; cependant, il ne s'agit pas d'une revue de la littérature sur l'évaluation ni d'un tour d'horizon complet du sujet. Le dossier est un document de référence: le professeur pourra facilement se référer, au moment opportun, au texte qui l'intéresse. De plus, le dossier est conçu de façon à pouvoir s'enrichir à chaque publication de textes complémentaires.

Le dossier est divisé en cinq sections. La section 1 regroupe les textes définissant les concepts d'évaluation. Plus particulièrement, le texte *Evaluer? Mesurer?* explique brièvement le vocabulaire de base: mesure, évaluation, évaluation formative, évaluation sommative, évaluation continue, évaluation normative et évaluation critériée. Le texte *Assurer à tous la compétence* définit le concept du «mastery learning», indique ses conséquences et certaines modalités d'application. Dans *Les modèles d'évaluation*, on explicite les modèles d'évaluation sommative-normative et formative-critériée en en donnant les caractéristiques, en identifiant certains problèmes et éléments de solution qui leur sont relatifs. Le texte *Il ne faut pas confondre l'autonotation et l'auto-évaluation* éclaire le concept de l'auto-évaluation et en donne certaines applications, tout en indiquant les dangers inhérents à l'autonotation.

On trouve, dans la section 2, des textes présentant les processus d'élaboration de deux instruments d'évaluation différents. Le premier texte porte sur les étapes permettant de rédiger, d'organiser des objectifs, puis de construire une grille d'évaluation de l'apprentissage dans la situation des stages; le second a pour objet la démarche pour évaluer l'enseignement.

A la section 3, on présente des textes se rapportant à des pratiques particulières d'évaluation de l'apprentissage. Dans *Les tests sur les prérequis et les pré-tests*, on identifie les buts de ces deux types de tests et diverses modalités d'application de ceux-ci. *Faut-il pénaliser le choix au hasard?* s'adresse aux utilisateurs de tests objectifs composés de questions à choix multiple, c'est-à-dire de questions présentant un choix de réponses. On y identifie le problème posé par le choix d'une réponse au hasard. On y

analyse diverses solutions, et plus particulièrement celle des formules de correction.

A la section 4 apparaissent les travaux sur la notation. Le texte *Les systèmes de notation*, basé sur une revue de la littérature, présente divers systèmes de notation utilisés en Amérique du Nord eet examine l'utilité des notes. Le texte *Abolir les notes, pourquoi pas?* expose divers problèmes suscités par la notation et son usage, pour ensuite explorer la possibilité d'abolir les notes.

La section 5 regroupe diverses expériences en évaluation. Le texte *Un système informatique de contrôle de l'apprentissage* présente une expérience d'enseignement individualisé de type modulaire dont le système d'évaluation s'appuie sur l'utilisation de l'ordinateur. Il s'agit d'un cours fondamental pour les ingénieurs, un cours de statique s'adressant à plus de 500 étudiants. *La banque d'objectifs et la banque d'items: un exemple d'application* résume la création et l'exploitation d'une banque d'items pour évaluer l'apprentissage d'étudiants en gynécologie-obstétrique. On examine dans *L'évaluation de la formation clinique en ergothérapie* les étapes de l'élaboration d'une grille d'évaluation formative des apprentissages faits lors des stages.

Des textes et de nouvelles sections sont dès maintenant prévisibles et pourront s'ajouter au fur et à mesure de leur production.

Vos commentaires sur ce dossier seront les bienvenus afin de nous aider à améliorer la prochaine édition.

F.F.

Section 1

La définition des concepts d'évaluation

Évaluer? Mesurer?

France Fontaine

Plusieurs termes prêtent à confusion dans le domaine de la mesure et de l'évaluation. Afin d'offrir un cadre de référence qui facilite les discussions, nous tenterons de définir les termes suivants:

- L'évaluation
- La mesure
- L'évaluation formative
- L'évaluation sommative
- L'évaluation critériée
- L'évaluation normative

Des exemples permettront de mieux distinguer ces concepts.

L'évaluation, la mesure

Evaluer, c'est porter un jugement de valeur. Dans nos institutions d'enseignement, l'objet de ce jugement de valeur peut être l'acquis d'un étudiant dans un tel cours, l'enseignement donné par tel professeur, le programme de telle faculté, etc.

La mesure est, quant à elle, non pas un jugement de valeur, mais *la cueillette et la classification systématique de données quantifiables sur un objet ou une classe d'objets*. La mesure consiste en l'application d'une règle d'attribution de symboles à des objets. Cependant, les mesures obtenues dans le domaine de l'enseignement sont toujours entachées d'une certaine part d'erreur; si celles dans le domaine des sciences dites exactes n'en sont pas exemptes, il est facile de s'imaginer que cela est d'autant plus probable dans le domaine des sciences humaines.

Les résultats d'une mesure peuvent être utilisés pour évaluer : nous interprétons alors les résultats de la mesure par rapport à un cadre de référence. Nous pouvons aussi évaluer sans mesurer, c'est-à-dire en nous guidant sur des préjugés, des opinions,...

Donnons quelques exemples afin d'illustrer ces concepts.

Lorsque vous dites, après avoir regardé votre thermomètre, qu'il fait 35°C, vous faites état d'une mesure. Après quelque temps, vous avez pris l'habitude d'utiliser cette nouvelle échelle de température et si, lorsque l'on vous dit qu'il va faire 35°C, vous répondez qu'il va faire très chaud, vous portez alors un jugement, c'est-à-dire que vous évaluez.

Un de vos étudiants a obtenu 75% à l'examen que vous avez administré. Ce pourcentage est la mesure fournie par l'instrument qu'est votre examen et il n'est ni bon ni mauvais en soi. Cependant, lorsque vous dites que cet étudiant a très bien réussi, vous évaluez ce résultat puisque vous lui donnez une signification.

Pourquoi évaluer?

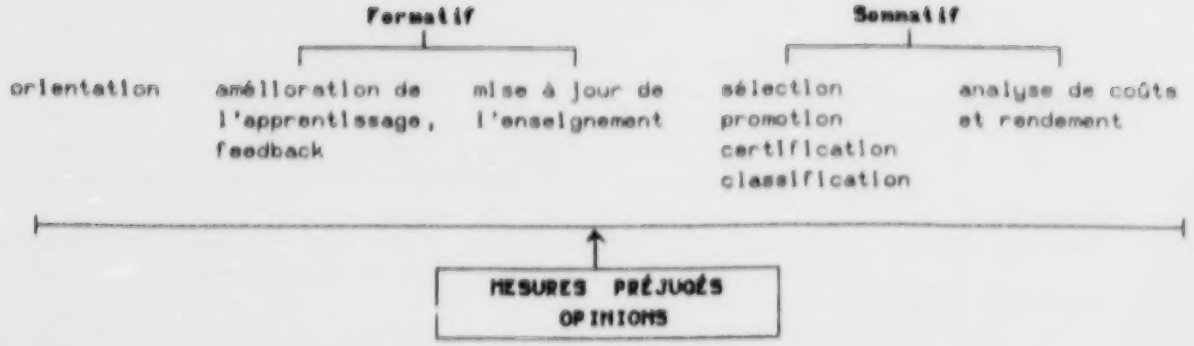
Le professeur doit se demander pourquoi il évalue plutôt que de s'interroger à savoir s'il doit mesurer ou évaluer. La question qu'il devrait se poser serait la suivante : en fonction de quel type de décision dois-je évaluer?

Nous distinguons cinq types principaux de décisions ou rôles de l'évaluation :

- les décisions qui visent à orienter les étudiants dans leur choix de cours et de programmes d'études ;
- les décisions qui visent à améliorer l'apprentissage des étudiants ;
- les décisions qui visent à faciliter la mise à jour de l'enseignement donné ;
- les décisions qui visent à sélectionner et à classifier ;
- les décisions qui visent à déterminer le coût et le rendement.

Nous regroupons sous le terme *évaluation sommative les évaluations faites à des fins de sélection et d'analyse de coûts et de rendement*. C'est le type d'évaluation qui intéresse plus particulièrement l'administration. Par contre, *l'évaluation faite dans le but d'améliorer l'apprentissage et l'enseignement s'appelle formative* et s'adresse au professeur et à l'étudiant.

STRUCTURE DE LA DÉCISION



Evaluation formative, évaluation sommative

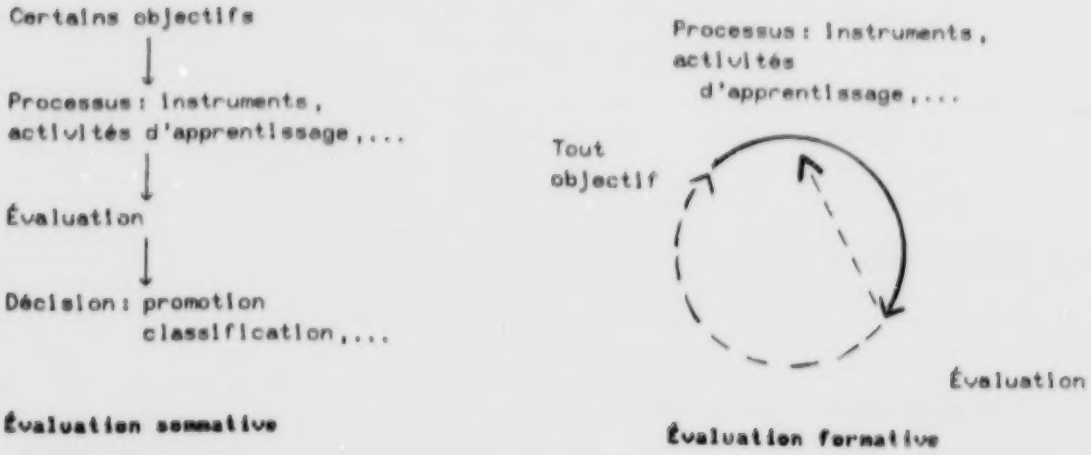
Nous entendons par *évaluation sommative un jugement définitif qui met un terme à une activité*. Un diplôme consacrant la fin d'un programme en est un exemple. Il s'agit d'évaluation sommative lorsque vous jugez qu'un étudiant a réussi son cours en obtenant la note de passage à un examen final unique ou s'il a totalisé 60% et plus dans les tests ou les travaux donnés durant la session. Nous parlons encore d'évaluation sommative lorsque l'Université juge que tel programme d'études est meilleur que tel autre et qu'il devrait, en conséquence, remplacer ce dernier.

L'évaluation *formative*, contrairement à l'évaluation sommative qui a pour but de certifier l'apprentissage, *vise plutôt à l'améliorer*. Elle a lieu pendant les activités d'apprentissage et permet un retour en arrière si des corrections s'imposent.

Voici un premier exemple d'évaluation formative: vous évaluez les acquis de vos étudiants afin de leur dire s'ils ont atteint les objectifs de votre cours, s'il y a lieu de reprendre telle activité d'apprentissage ou d'utiliser un autre moyen pour corriger les faiblesses ainsi dépistées. Vous faites une évaluation formative lorsque vous demandez à un étudiant de reprendre telle ou telle partie d'un travail qui ne correspond pas aux critères préalablement établis et qu'il y a échange avec l'étudiant afin de lui permettre d'apporter les correctifs qui s'imposent. C'est encore de ce type d'évaluation qu'il s'agit lorsque vous cherchez à améliorer votre cours en révisant les objectifs de celui-ci, les moyens que vous avez pris pour les atteindre et les évaluer. Cela vous conduit alors à un système non-linéaire, fondamentalement différent de l'évaluation sommative, puisqu'il y a toujours retour aux diverses composantes du système et une remise en question de celles-ci.

Il ne faut pas confondre l'évaluation dite «continue» et l'évaluation formative. L'évaluation «continue» est une évaluation sommative cumulative: elle ne permet pas de retour en arrière.

L'ÉVALUATION SOMMATIVE ET L'ÉVALUATION FORMATIVE



Cadre de référence

Nous distinguons habituellement deux cadres de référence ou deux façons de porter un jugement de valeur. Si l'évaluation *vise à situer la position d'un objet dans sa catégorie*, en ce qui concerne la caractéristique envisagée, il s'agit alors d'une *évaluation normative*. Par exemple, on compare le rendement d'un étudiant à celui des autres étudiants de son groupe; ces étudiants peuvent être ordonnés à partir de leurs notes brutes à un examen, de leurs rangs centiles, etc... Les analyses d'examen qui permettent un rejet des questions non-discriminatoires découlent de l'évaluation normative de même que les méthodes de transformation de résultats en stanines, à partir de la courbe normale. L'évaluation est dite *critériée* lorsqu'elle *vise à déterminer si l'objet est satisfaisant par rapport à un critère, indépendamment de sa position dans sa classe*. L'objet n'est pas comparé aux autres objets de sa classe mais à un critère. Par exemple, on se demande si un étudiant donné a atteint tel objectif d'apprentissage, s'il est apte à un apprentissage subséquent et ce, indépendamment de la réussite des autres étudiants de son groupe.

Donnons d'autres exemples, afin de mieux illustrer les différences entre ces deux cadres de référence. C'est d'évaluation normative qu'il est question lors des compétitions olympiques en natation: le gagnant de la médaille d'or est le meilleur parmi les inscrits à une compétition. Cependant, l'obtention de l'insigne du niveau juvénile relève plutôt de l'évaluation critériée puisqu'elle n'est pas liée à la performance du groupe, mais à la performance de l'individu, par rapport à des critères tel que de nager une certaine distance en un laps de temps donné, etc. Qu'un individu soit le meilleur parmi un groupe d'individus ne sachant pas nager ne lui permettra pas d'obtenir cet insigne.

Soulignons que l'évaluation normative est plus souvent liée à l'évaluation sommative, tandis que l'évaluation formative est plus souvent associée à l'évaluation critériée. L'évaluation formative ne cherche pas à reproduire la courbe normale mais vise plutôt à amener le plus d'étudiants à acquérir certaines compétences. Cependant, les gens dans nos institutions ont rarement recours à des formes d'évaluation qui soient entièrement normatives ou entièrement critériées. Nous retrouvons plutôt des pratiques «mixtes» qui prennent le nom du type d'évaluation dominant, soit critérié, soit normatif.

Notons enfin qu'il faut savoir pourquoi nous voulons évaluer afin de choisir le type d'évaluation et le cadre de référence qui conviennent.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogiq..., Université de Montréal, Vol. 1 No 3, Sept. 1978, pp.8-10.

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Assurer à tous la compétence

Claire Felx-Larivière
France Fontaine

Etes-vous las de constater qu'un petit nombre seulement de vos étudiants réussissent bien votre cours? Qu'un certain nombre obtiennent la note de passage même si vous estimez leur rendement plutôt médiocre alors que d'autres échouent?

Pour corriger cette situation, un concept global est apparu, le «mastery learning[1]», qui définit une nouvelle approche éducative.

LE «MASTERY LEARNING»

La compétence ne doit plus être le privilège d'une minorité mais doit être accessible à tous. Le «mastery learning» repose sur les postulats voulant que ce qui différencie les étudiants, c'est leur rythme personnel d'apprentissage et que la plupart (sans doute plus de 90%) sont capables de réussir un cours donné. On peut définir ce concept comme une pratique qui vise à fournir à tous les étudiants, à l'aide de moyens variés, la possibilité d'acquérir la compétence. Cela suppose des apprentissages identifiables et circonscrits qu'un programme veut faire acquérir à tous les étudiants, un certain accord sur les standards de compétence et des moyens pour déterminer si ces standards ont été atteints.

Le «mastery learning» est d'abord lié à l'apprentissage, mais il est aussi en étroite relation avec des concepts d'évaluation. Il suppose des buts, des objectifs, des standards et une évaluation critériée[2].

Il faut, aussi et surtout, que l'on adapte l'apprentissage à l'étudiant par une variété de moyens et de méthodes[3]. L'évaluation formative devient nécessaire puisqu'un feedback rapide permet un meilleur apprentissage et que l'étudiant peut progresser à son rythme. C'est à l'enseignement de faire en sorte, par les moyens d'apprentissage et de feedback, que tous, ou presque, réussissent.

Certains exemples d'utilisation:

Des professeurs ont établi leurs standards en s'inspirant de ce que leurs meilleurs étudiants réussissaient à atteindre au cours des années précédentes. Des tests critériés et formatifs ont été élaborés afin de dire à chaque étudiant, selon sa progression dans le cours, s'il a réussi ou non à atteindre le standard indiqué pour tel objectif. Dans le cas où le standard n'est pas atteint, un diagnostic lui permet de se reprendre tout en sachant ce qui ne va pas.

Dans d'autres cas, les professeurs ont établi des standards minimaux de succès. Il peut s'agir de préciser ce qu'ils jugent essentiel à la poursuite de l'apprentissage. D'un étudiant en génie civil, on peut exiger un minimum de connaissances (mathématiques, résistance des matériaux, etc.) lui permettant de préparer les plans d'un pont solide.

Les conséquences

Parmi les conséquences du «mastery learning», on constate que les instruments de mesure utilisés ne reproduisent pas une courbe normale ou quelque chose qui s'en approche. Ceci n'a rien de surprenant, car c'est en réaction contre l'application de la courbe normale dans le monde de l'éducation que le «mastery learning» a été conçu. Si l'on croit qu'un petit nombre d'étudiants sont forts, qu'un certain nombre ont un rendement moyen et que les autres sont faibles, on enseignera et on évaluera par des mesures en conséquence. Or, l'enseignement dans ce contexte ne cherche pas à étaler les étudiants sur un continuum en les différenciant le plus possible, mais il vise à fournir aux étudiants les moyens d'atteindre les objectifs du cours. On utilise donc des mesures critériées et l'on obtient une distribution des résultats fort différente: ceux qui répondent aux critères et ceux qui n'y répondent pas. Les exigences ne varient plus selon le groupe, puisque les standards sont les mêmes, qu'il s'agisse d'un groupe fort ou d'un groupe faible.

L'usage de la courbe normale est de plus en plus contesté dans le monde de l'éducation, que ce soit pour évaluer l'apprentissage de l'étudiant ou pour évaluer l'enseignement dispensé. Si l'on peut encore émettre l'hypothèse que la facilité pour les mathématiques, par exemple, est distribuée dans la population selon cette courbe, il faut tenir compte du fait que, au niveau universitaire, la clientèle est fort restreinte et n'est guère représentative de toute la population. De plus, ce que l'on cherche à mesurer en classe, ce ne sont pas les aptitudes mais le rendement, s'il y a apprentissage ou non. Un professeur de calcul différentiel ne cherche pas à vérifier la plus ou moins grande facilité qu'ont ses étudiants à faire des mathématiques mais il veut

savoir si chacun de ses étudiants peut effectuer tel type de dérivée. Une étude démontre même que la corrélation entre l'aptitude et rendement est presque inexistante lorsque l'on a su assurer la compétence à la majorité des étudiants, en respectant leur rythme d'apprentissage et en utilisant divers moyens d'enseignement.

Le jeu de la moyenne est contrecarré par cette pratique. La facilité dans un domaine ne peut plus compenser des lacunes dans un autre. Par exemple, il n'est plus possible que la dextérité manuelle, en chirurgie dentaire, soit contre-balancée par une capacité de jugement, ou qu'une carence, par rapport à un niveau minimal de synthèse, soit rachetée par une bonne capacité d'analyse. Il ne peut plus se produire que des étudiants se désintéressent de leur cours, parce qu'ils ont déjà accumulé suffisamment de points pour s'assurer la note de passage, alors que le professeur juge qu'il reste des choses importantes à apprendre. Un étudiant qui manifeste des faiblesses au chapitre 3, où l'on trouve des notions fondamentales pour les cours à venir, ne peut compenser ces faiblesses par de bons résultats aux tests portant sur les autres chapitres. On peut même entrevoir la possibilité de poursuivre des objectifs, jusqu'à présent laissés pour compte, - comme l'auto-évaluation et l'auto-formation -, qui s'inséraient mal dans un calcul de moyenne.

Cette pratique a aussi pour effet de diminuer la compétition entre les étudiants et d'améliorer la perception qu'un étudiant a de lui-même. Identifié, dès l'école primaire, comme bon, moyen ou faible, l'étudiant se conforme souvent à cette étiquette et se perçoit comme tel.

Les problèmes

Il faut évidemment beaucoup de temps pour planifier un tel cours, préparer les tests critériés, concevoir des moyens d'apprentissage variés et un système d'évaluation formative. Il est difficile de préciser les apprentissages que l'on veut favoriser chez tous les étudiants, de bien choisir les standards de compétence et de produire les tests permettant de vérifier si l'étudiant les a atteints ou non. Evidemment, lorsque tout cela est sur pied, la tâche du professeur est facilitée car il subsiste moins d'ambiguités.

Le fait de vouloir respecter le rythme de l'étudiant pose certains problèmes dans notre système éducatif. Trop souvent, on ne peut permettre des rythmes variés qu'à l'intérieur de quinze ou seize semaines de cours. Cette contrainte peut aller jusqu'à compromettre le respect de ces rythmes variés. Des professeurs contournent la difficulté en inscrivant la mention «remise» au dossier. Ceci permet à l'étudiant, en cas de nécessité, de compléter son cours en dehors du cadre restreint d'une session académique. Certains espèrent voir s'instaurer le respect des rythmes d'apprentissage à tout un programme, ce qui entraînerait alors une modification du système.

Il faut aussi souligner que des enseignants se sont servis de ce modèle en précisant le minimum requis. Malheureusement, certains ont limité l'apprentissage à ce minimum. Il faut plutôt fournir à l'étudiant les ressources nécessaires et l'encourager à poursuivre plus loin sa formation.

Il est sans doute plus difficile de faire participer à chaque session les étudiants au choix des objectifs et des critères de compétence, étant donné le matériel à préparer. Sauf dans les cas où l'étudiant choisit ses objectifs dans une banque, c'est presque toujours le professeur ou un groupe de professeurs qui les déterminent. La participation étudiante est donc souvent inexistante.

On adresse souvent à notre société nord-américaine le reproche de rechercher avant tout l'efficacité. Ce reproche s'étend sans doute à cette volonté de rentabiliser l'enseignement.

Si certaines personnes pensent que ce modèle peut s'appliquer à tout enseignement, d'autres le réservent à certains apprentissages, alléguant que ce modèle s'appuie sur un objectif final plutôt que sur une continuité de l'apprentissage. La capacité de synthèse, par exemple, ne requiert-elle pas un apprentissage échelonné sur toute une vie? À tout le moins, peut-on dire qu'il est plus difficile, mais peut-être possible, dans ces cas-là, de préciser des étapes, de distinguer des niveaux d'apprentissage. La compétence n'est pas prise comme un tout: on ne définit pas globalement la compétence dans telle science ou dans tel art. Il s'agit plutôt d'une compétence morcelée devant des apprentissages précis.

Conclusion

En résumé, on peut dire que, pour assurer à tous la compétence, on utilise des moyens variés d'apprentissage, dont l'évaluation formative qui permet de respecter le rythme de chacun des étudiants. Une évaluation de type critérié est de mise pour vérifier l'atteinte des standards, s'il y a compétence ou non. L'imagination des professeurs a donné naissance à diverses modalités d'application du «mastery learning», par exemple, des cours modularisés, individualisés, etc. Cependant, il reste encore à explorer d'autres modalités d'application.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 2 No 3, Sept. 1977, pp. 17-19.

1. Le texte se réfère au concept appelé en anglais «mastery learning» ou «learning for mastery». Pour un supplément d'information, vous pouvez vous référer au chapitre 3 du volume suivant: Bloom, Benjamin S., Hastings, J. Thomas, Madaus, George F., *Handbook on Formative and Summative Evaluation of Student Learning*. New-York, McGraw-Hill Inc., 1971, 923 p..

2. Vous trouverez l'explication de ces termes dans le texte suivant: Fontaine, France, «Évaluer? Mesurer?» dans *Dossier sur l'évaluation*, pp. 5-9.

3. L'étude des styles d'apprentissage peut faciliter cette tâche. Vous pouvez consulter à ce sujet: Scholer, Marc et al., *Profil d'apprentissage et enseignement personnalisé d'après Joseph E. Hill*. Montréal, Service général des moyens d'enseignement, ministère de l'Education, 1974, 139 p., (code DTE-01-0474).

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Les modèles d'évaluation de l'apprentissage

France Fontaine

Les termes de base tels que: évaluation formative, évaluation sommative, évaluation critériée et évaluation normative ont déjà été définis dans le texte *Évaluer? Mesurer?* Le présent texte se veut donc une suite explicitant les modèles d'évaluation dans le contexte plus restreint de l'apprentissage de l'étudiant.

L'évaluation sommative est habituellement liée à un cadre de référence normatif, tandis que l'évaluation formative est plus souvent pratiquée dans un cadre de référence critérié. On retrouvera donc, pour l'évaluation sommative et normative, puis pour l'évaluation formative et critériée, les caractéristiques, problèmes et éléments de solution afférents à ces modèles. On exposera ensuite brièvement les possibilités de modèles mixtes.

L'ÉVALUATION SOMMATIVE ET NORMATIVE

Rappelons qu'on a défini *l'évaluation sommative* comme étant un jugement définitif qui met un terme à une activité. L'évaluation a donc pour but de déterminer si un étudiant a réussi ou non, à des fins de certification, de promotion et ainsi de suite. Elle se situe à la fin de la période consacrée à l'enseignement-apprentissage, c'est-à-dire à la fin du module, du cours, du programme, alors qu'un retour en arrière n'est plus possible.

Lorsqu'on utilise *un cadre de référence normatif* pour poser un jugement, c'est-à-dire que l'on vise à situer la position d'un objet dans sa catégorie (en ce qui concerne la caractéristique envisagée), on identifie le degré de réussite d'un étudiant par des comparaisons avec ses pairs. On se réfère à des normes.

Caractéristiques

a) L'évaluation sommative ayant un caractère définitif, c'est-à-dire une grande importance pour l'avenir de l'étudiant, elle présente les caractéristiques suivantes:

- elle est *valide* c'est-à-dire qu'elle doit mesurer tout ce qu'il faut mesurer et cela seulement;
- elle est *fidèle*, c'est-à-dire que les mesures doivent être précises et stables;
- elle se doit d'être *objective* c'est-à-dire d'assurer un traitement égal pour tous;
- elle est nécessairement *formelle*, c'est-à-dire que les règles d'évaluation doivent être clairement établies et observées;
- elle ne porte habituellement que sur un *échantillon des objectifs* du cours, c'est-à-dire sur les objectifs jugés les plus importants;
- elle n'a que *peu de conséquences sur la planification* du cours.

b) Comme l'évaluation normative vise à situer les étudiants les uns par rapport aux autres selon leur apprentissage, il faut choisir des instruments de mesure qui facilitent les comparaisons entre les étudiants. Ainsi, pour un test écrit, on choisira les questions ou items qui discriminent le plus, qui distinguent le plus les étudiants les uns des autres. C'est pourquoi on éliminera d'un test les questions dites faciles, parce que répondues par presque tous les étudiants, ou celles jugées trop difficiles, parce que répondues par un très petit nombre d'étudiants.

Problèmes

1° Souvent, les instruments de mesure (examens, tests, etc.) utilisés ne sont guère valides, fidèles et objectifs.

2° On analyse rarement les instruments de mesure à partir des réponses des étudiants, par manque de temps, de moyens ou de motivation.

3° Les points sur lesquels porte l'évaluation constituent parfois un échantillon plus ou moins représentatif de l'ensemble des objectifs du cours: on n'évalue pas les objectifs essentiels.

4° Il arrive qu'on accorde une importance absolue à des instruments de mesure, qu'ils soient bien construits ou non. Ceci mène à l'automatisme dans la prise de décision: par exemple, lorsque l'étudiant obtient 60% et plus à son test, il réussit, sinon il échoue.

5° Des comportements peu souhaitables sont fréquemment liés à l'évaluation sommative. L'étudiant peut travailler en fonction de sa note finale; il peut chercher à deviner ce qui est important pour le professeur si celui-ci ne l'a pas suffisamment explicité; il arrive souvent qu'il y ait bourrage de crâne à la veille de l'examen.

6° Une trop grande anxiété de la part de certains étudiants en période d'examens diminue la fiabilité des résultats.

7° Comme on identifie les difficultés d'apprentissage des étudiants qu'à la fin de la période d'enseignement-apprentissage, il est trop tard pour apporter les correctifs qui s'imposent. Il peut donc arriver qu'un étudiant poursuive ses études malgré certaines lacunes ou qu'on l'oblige à recommencer tout son cours même si ses lacunes sont concentrées dans telle ou telle partie du cours.

3° L'ambiguïté est parfois grande entre le rôle du professeur comme facilitateur de l'apprentissage et son rôle de juge de ces apprentissages.

9° L'étudiant, n'ayant été contrôlé qu'à la fin de l'apprentissage, n'a guère pu développer de critères pour s'auto-évaluer[1] correctement. Une des conséquences en est que plusieurs étudiants considèrent injustifiés leurs résultats scolaires.

10° Lorsqu'on évalue suivant le cadre de référence normatif, l'échec ou le degré de réussite d'un étudiant dépend de la distribution des résultats de l'ensemble des étudiants. Ainsi, selon que le groupe auquel on le compare est plus ou moins fort, il pourra, à la limite, échouer ou réussir.

Éléments de solution

Soulignons d'abord que si certains problèmes peuvent être résolus, partiellement ou complètement, certains autres sont inhérents à cette forme d'évaluation (par exemple, les problèmes 7 et 9). Il faut s'attendre à ce que ces problèmes subsistent si l'évaluation n'a pas pour but de les résoudre.

- Afin d'améliorer la construction des instruments de mesure, on pourrait mieux identifier les objectifs du cours, les critères de réussite et distinguer entre les objectifs terminaux[2] et les objectifs intermédiaires. Comme l'évaluation sommative se limite habituellement à une partie des objectifs, il est possible de ne la faire porter que sur les objectifs jugés importants, par exemple les objectifs terminaux.

  On peut aussi classer les objectifs selon les niveaux d'une taxonomie donnée (par exemple connaissance, compréhension, application, analyse, synthèse et évaluation pour le domaine cognitif, selon la taxonomie de Bloom[3] afin de préparer un tableau de spécification.

  Un tableau de spécification est un tableau à double entrée: une pour le contenu (ou objectif), une autre pour le niveau taxonomique. On indique alors la pondération, par exemple en pourcentage, qu'il faut accorder dans l'instrument de mesure à tel contenu traité à tel niveau taxonomique. Donc, si dans un cours donné, la compréhension de la notion x a été jugée comme devant recevoir le poids de 10% sur le total de 100% pour le cours ou la partie du cours à évaluer, l'instrument de mesure devra, par le nombre de questions sur la compréhension de cette notion et par la pondération accordée à ces questions, refléter ce 10%.

  Représentons ici un schéma de tableau de spécification pour une partie de cours comportant onze notions classées suivant la taxonomie de Bloom du domaine cognitif.

NOTION (ou objectif) / NIVEAU TAXONOMIQUE	CONNAISSANCE	COMPRÉHENSION	APPLICATION	ANALYSE	SYNTHÈSE	ÉVALUATION	TOTAL
x_1	8%	-	-	-	-	-	8%
x_2	-	-	4%	-	-	-	4%
x_3	-	4%	4%	4%	-	-	12%
x_4	2%	2%	2%	4%	-	-	10%
x_5	-	-	4%	-	4%	-	8%
x_6	-	-	-	-	4%	-	4%
x_7	2%	2%	-	-	-	-	4%
x_8	-	-	-	10%	6%	-	16%
x_9	-	-	2%	6%	-	-	8%
x_{10}	-	-	6%	-	8%	-	14%
x_{11}	-	4%	6%	2%	-	-	12%
TOTAL	12%	12%	28%	26%	22%	-	100%

- Dans cet exemple, la notion x_1 étant présentée pour la première fois, il est normal que l'on ne vise que le niveau de connaissance. Les notions x_2, x_5, x_6, x_8, x_9 et x_{10} ayant été évaluées au niveau de la connaissance et de la compréhension par des examens antérieurs, on peut donc les évaluer à des niveaux cognitifs supérieurs: il est normal de chercher à atteindre de tels niveaux dans l'enseignement universitaire. Il est à noter que le niveau «évaluation» n'est pas visé: il n'est pas nécessaire d'évaluer tous les niveaux taxonomiques dans un cours.

  Cette façon de procéder à l'aide d'un tableau de spécification permet d'éviter les questions portant sur des détails peu importants de la matière. Il faut cependant veiller à ce que les questions correspondent bien à l'objectif énoncé et visent le bon niveau taxonomique.
- Il est de plus en plus facile d'analyser les instruments de mesure (surtout composés de questions à choix multiple, c'est-à-dire de questions offrant un choix de réponses) car le nombre de programmes de traitement informatisés créés à cette fin est sans cesse grandissant. On peut ainsi accumuler les statistiques concernant les questions, éliminer ou reformuler les questions mal construites et le reste.
- Il faut s'assurer qu'il y a accord sur la réponse à donner à une question: trop souvent il y a divergence chez les professeurs. C'est l'étudiant qui en paye les frais.

- On peut augmenter l'objectivité dans la correction de tests ou de travaux en préparant des instruments tels qu'une liste des bonnes réponses avec le nombre de points à accorder pour chacune, une liste de critères d'évaluation détaillés et pondérés et le reste.
- Lorsqu'on assure aux étudiants de bonnes conditions d'administration des instruments de mesure (instructions claires, environnement favorable, etc.), on contribue à augmenter la fidélité de ces instruments. Dans le cas de questions à choix multiple, on peut diminuer les erreurs de mesure dues au choix d'une réponse au hasard en accordant suffisamment de temps aux étudiants pour répondre, en évitant les questions trop difficiles, en posant d'abord les questions faciles puis les questions difficiles et ainsi de suite.
- Comme tout instrument d'évaluation de l'apprentissage ne mesure pas directement l'objet à évaluer, il y a toujours erreur de mesure. Donc le résultat observé d'un étudiant n'est pas nécessairement son résultat véritable. On peut cependant estimer cette erreur de mesure à partir du coefficient de fidélité du test. On peut ainsi calculer, avec un certain degré de probabilité, un intervalle de confiance autour du résultat observé, dans lequel se situerait le résultat véritable de l'étudiant. On peut, de la sorte, diminuer l'arbitraire de la note de passage en appliquant ce calcul d'intervalle de confiance à la note de passage.
- Si le professeur communique à l'étudiant ses intentions pédagogiques par un énoncé explicite des objectifs du cours, en indiquant si certains objectifs sont prioritaires et si les instruments de mesure sont conformes aux objectifs, l'étudiant n'aura pas à deviner ce qui est important pour le professeur. Il pourra se concentrer sur les apprentissages à faire et il est probable qu'il sera moins anxieux par rapport à l'évaluation.
- Il est parfois possible d'éviter le bourrage de crâne à la veille des examens en vérifiant si les objectifs que l'on vise sont ou non des objectifs de connaissance. Lorsqu'il ne s'agit pas de tels objectifs, on peut en informer les étudiants et construire les instruments de mesure en conséquence. Il est peut-être possible aussi de diminuer la tendance au bachotage par un plus grand nombre d'évaluations sommatives, c'est-à-dire par l'évaluation continue[4].
- Comme il n'est pas obligatoire que l'évaluation sommative soit également normative, on peut se demander si le cadre de référence normatif est bien celui qui convient, si on en tire bien les renseignements désirés. Serait-il plus avantageux de situer l'apprentissage de l'étudiant par rapport aux objectifs ou est-il plus utile de comparer l'étudiant aux autres étudiants de son groupe?

L'ÉVALUATION FORMATIVE ET CRITÉRIÉE

L'évaluation formative vise à améliorer l'apprentissage: elle a lieu pendant l'apprentissage et permet un retour en arrière si des corrections s'imposent. Ce qui la distingue de l'évaluation sommative, c'est qu'une intervention est encore possible; ce n'est pas un jugement définitif. Elle est en elle-même une activité d'apprentissage, l'étudiant apprenant par cette évaluation, qui le renseigne sur ses succès et échecs.

Lorsqu'un *cadre de référence critérié* est utilisé pour poser un jugement, on cherche à déterminer si l'objet est satisfaisant par rapport à un critère, indépendamment de sa position dans sa classe.

On se demande si l'étudiant atteint les critères prédéterminés, et ce, sans référence aux autres étudiants de son groupe. On se réfère donc à des critères plutôt qu'à des normes.

Caractéristiques

a) Comme l'évaluation formative a pour but d'aider l'étudiant dans son apprentissage, elle présente les caractéristiques et entraîne les conséquences suivantes:

- Elle sera suffisamment fréquente pour que l'étudiant ait le temps de revenir sur des apprentissages, s'il y a lieu. Ceci lui permettra également de ne pas s'égarer trop longtemps sur une fausse piste, comme l'étudiant qui ne comprendrait pas les instructions données par le professeur pour un travail ou qui procéderait de façon erronée pour résoudre un certain type de problèmes. Qu'il s'agisse de renforcement ou de corrections à apporter par d'autres activités d'apprentissage, il faut tarder le moins possible à informer l'étudiant de ses succès et de ses erreurs. Il ne faut cependant pas confondre cette évaluation avec une forme d'évaluation qu'on peut qualifier de cumulative sommative, souvent appelée continue[5], qui ne permet pas de revenir sur des apprentissages.
- Elle exige que les résultats de l'évaluation soient connus rapidement des étudiants, pour les mêmes raisons que celles énoncées au sujet de la fréquence de l'évaluation. Il n'est pas rare de voir les résultats d'une évaluation sommative remis plusieurs jours, voire plusieurs semaines, après les épreuves d'évaluation. Une telle pratique n'aide guère l'étudiant à combler les lacunes de ses apprentissages, surtout dans le cas où il est limité à une période déterminée (souvent quinze semaines) pour atteindre les objectifs du cours.
- Elle nécessite la prescription d'activités d'apprentissage si l'objectif n'est pas atteint. Il peut s'agir de reprendre certaines des mêmes activités d'apprentissage ou encore de procéder à de nouvelles activités d'apprentissage mieux adaptées aux besoins de l'étudiant (par exemple: visionner un film au lieu des lectures prescrites). Le professeur doit penser à ces autres activités d'apprentissage et, s'il y a lieu, concevoir et réaliser le matériel qu'elles supposent.

- Elle est transparente, c'est-à-dire que les étudiants doivent savoir ce qu'on attend d'eux. Il n'y a pas de secret quant à l'objet de l'évaluation.
- Elle n'est pas nécessairement formelle. Par exemple, elle peut être basée sur des instruments d'auto-évaluation[6] mis à la disposition des étudiants avec prescription d'activités d'apprentissage. Ou encore le professeur peut aider quelques étudiants dont il a observé les difficultés d'apprentissage et revérifier ensuite les acquis.
- Elle exige assez souvent une planification de cours plus poussée que l'évaluation sommative. C'est habituellement le cas lorsque les étudiants sont nombreux ou lorsqu'on veut bien intégrer l'évaluation formative à l'enseignement, par exemple en rendant plus formelle l'évaluation formative. Il faut choisir à quel moment auront lieu les contrôles, prévoir quand et comment seront revérifiés les apprentissages des étudiants qui ont manifesté des lacunes au premier contrôle et le reste.
- Comme l'incidence de l'évaluation formative sur l'étudiant n'a pas de caractère définitif, contrairement à l'évaluation sommative, plusieurs auteurs sont d'avis qu'il n'est pas nécessaire que les instruments d'évaluation soient aussi valides et fidèles que ceux qui sont requis par l'évaluation sommative.
- Afin de dépister les problèmes d'apprentissage, elle demande que tout objectif soit vérifié, qu'il s'agisse d'un objectif intermédiaire, d'un objectif jugé facile à atteindre, etc. Le rôle de l'évaluation formative peut être des plus importants lorsqu'il y a des séquences d'apprentissage.
- Elle conduit assez souvent à l'individualisation de l'enseignement[7], au respect du rythme d'apprentissage. On permet à l'étudiant de reprendre des apprentissages sans le contraindre à le faire en dehors des heures de cours, sans le soumettre au rythme du groupe. Dans ce cas, le rôle du professeur est celui d'un facilitateur de l'apprentissage plutôt que d'un dispensateur de connaissances.

b) Comme l'évaluation critériée vise non pas à situer l'étudiant par rapport aux autres étudiants, mais à dire si l'étudiant atteint ou non les critères prédéterminés:

- on n'éliminera pas des instruments les questions trop faciles ou trop difficiles; il suffit qu'elles renseignent sur des apprentissages pertinents;
- elle se fera habituellement en fonction d'objectifs spécifiques[8] et on dira à l'étudiant qu'il a réussi ou non à atteindre tel objectif.

L'évaluation formative et critériée est souvent reliée à ce qu'on appelle le «mastery learning[9]». On cherche à fournir aux étudiants, à l'aide de moyens variés, la possibilité pour tous d'acquérir la compétence. L'évaluation formative est utilisée comme moyen d'apprentissage pour atteindre les standards, et les instruments sont critériés pour vérifier l'atteinte de ces standards. Comme on est plus exigeant en imposant des critères de réussite non pas sur l'ensemble du cours mais pour chacun des objectifs, on compense ces exigences par diverses facilités offertes à l'étudiant, comme l'évaluation formative.

Problèmes

Les deux premiers problèmes mentionnés pour l'évaluation sommative et normative (portant sur la validité, la fidélité, l'objectivité et l'analyse des instruments), peuvent aussi se présenter ici, bien que, comme on l'a mentionné dans les caractéristiques de l'évaluation formative, ces problèmes ont peut-être moins d'acuité. Fréquemment, certains des moyens envisagés pour résoudre ces deux problèmes font déjà partie de la démarche d'évaluation formative et critériée, par exemple rédiger des objectifs précis pour construire des instruments qui correspondent à ceux-ci.

De plus, on rencontre les difficultés suivantes:

1° La *tâche du professeur* est assez souvent accrue. (Parfois il doit consacrer plus de temps à la planification du cours, à la construction d'instruments d'évaluation, à la conception d'un système de correction ou, lors du déroulement du cours, au choix d'activités d'apprentissage différentes, à la correction, etc.)

2° L'*attitude du professeur* est un des facteurs qui déterminent le succès ou l'échec d'un processus d'évaluation formative dans lequel le professeur joue un rôle actif. L'étudiant doit sentir que son professeur désire l'aider à atteindre les objectifs du cours. Mais, tous les professeurs ne sont pas également à l'aise dans ce rôle.

3° Les étudiants éprouvent aussi des difficultés à modifier leurs *attitudes par rapport à l'évaluation et au professeur*. Par exemple, plusieurs croient qu'il leur est avantageux de cacher au professeur leurs difficultés d'apprentissage. (Lorsque l'évaluation formative est jumelée avec le «mastery learning[10]» et l'individualisation de l'enseignement[11], il n'est pas rare qu'on ait à faire face à des étudiants insécures qui vivent peut-être ce système pour la première fois).

4° Comme la *durée permise pour les apprentissages* d'un cours est généralement prédéterminée (habituellement, on dispose d'une quinzaine de semaines), on limite l'étudiant dans ses possibilités de faire des retours sur ses apprentissages. Autrement dit, le plein épanouissement de l'évaluation formative n'est pas facile à réaliser dans notre système éducatif.

5° Un minimum de *planification* est requis lorsque l'évaluation formative est intégrée à un cours. Ce minimum augmente au fur et à mesure qu'on rend plus formelle l'évaluation formative, qu'on s'adresse à un plus grand nombre d'étudiants.

Par exemple, il faut alors assurer une certaine souplesse dans le cours afin de permettre à l'étudiant de faire les réajustements nécessaires. Il faut choisir la fréquence des contrôles d'évaluation formative de telle sorte que l'étudiant reçoive fréquemment de l'information sur ses apprentissages sans pour cela que l'évaluation soit le seul moyen d'apprentissage, ou presque. L'équilibre n'est pas toujours facile à trouver. Il faut aussi préciser les modalités de recontrôle des étudiants qui ont manifesté des lacunes.

Il faut prendre ces décisions de sorte que le tout soit cohérent. Le «design» d'un tel cours peut être difficile à réaliser: le choix de mauvaises modalités d'application peuvent rendre nuls les avantages de l'évaluation formative.

6° Il n'est pas nécessairement facile d'assurer des *moyens d'évaluation équivalents* pour évaluer les étudiants ayant manifesté des lacunes dans leurs apprentissages, lors d'un contrôle précédent.

Éléments de solution

Comme pour l'évaluation sommative, il ne s'agit ici que de solutions partielles à des problèmes parfois complexes.

- Lorsque plusieurs professeurs peuvent travailler à un même cours, il est plus facile de mettre sur pied un système d'évaluation formative. Ils peuvent même assurer le cours en équipe. Dans le cas d'un professeur travaillant seul, celui-ci pourra planifier et expérimenter une partie du cours seulement, ou intégrer peu à peu les éléments d'un système d'évaluation formative de plus en plus complet tout en s'assurant de la cohérence à chacune des étapes. Ceci lui permettra d'identifier et de résoudre les problèmes au fur et à mesure.

  Un professeur peut aussi chercher à alléger sa tâche de correction par des auto-tests, par la mise sur pied d'un système de correction automatisée, par l'utilisation de questions à réponses brèves, etc., pour autant que ces moyens soient compatibles avec les objectifs du cours.
- Le professeur devrait se demander, même s'il est convaincu des avantages de l'évaluation formative, s'il accepte réellement le rôle d'aide à l'étudiant et si ses comportements face à celui-ci lui permettent de jouer efficacement ce rôle. Ainsi, un professeur qui semble manifester un certain mépris pour ses étudiants ne pourra pas vraiment réussir dans ce mode d'évaluation même s'il s'agit d'une fausse perception par les étudiants. Il lui faudra d'abord modifier les comportements qui sous-tendent cette perception étudiante (par exemple, éviter de ridiculiser ou de rabrouer un étudiant qui éprouve des difficultés d'apprentissage).
- Le professeur doit expliquer clairement aux étudiants les buts de l'évaluation formative et les inciter, par son attitude, par sa disponibilité, par les modalités du cours, à parfaire leurs apprentissages. Il doit faire croître, chez les étudiants, le sentiment de responsabilité face à leurs apprentissages et au contrôle de ceux-ci. Mais ce n'est pas facile. (Lorsqu'on a jumelé l'évaluation formative avec «le mastery learning[12]» et l'individualisation de l'enseignement[13], on peut rassurer l'étudiant en lui faisant sentir qu'il est suffisamment encadré, en lui exposant les résultats de cours similaires, etc.)
- Fréquemment, l'administration n'accorde que les quinze semaines du semestre et la période des examens pour compléter les dossiers étudiants. Il est parfois possible d'inscrire la mention «incomplet» ou «remise», accordant ainsi un délai à l'étudiant. Il peut alors terminer les apprentissages prévus entre les deux sessions ou à la session suivante, sans être obligé de recommencer les apprentissages correspondants aux objectifs déjà atteints.

- Avant d'annoncer aux étudiants un système d'évaluation formative, le professeur doit avoir estimé les ressources disponibles, prévu le déroulement du cours, préparé son matériel (par exemple: objectifs, matériel d'apprentissage, instruments d'évaluation, dossier étudiant), précisé les modalités d'application et le reste. C'est donc dire que le professeur doit s'assurer d'une certaine disponibilité auprès des étudiants durant le cours.
- La réalisation de banques de questions ou d'items[14] et les techniques d'analyse d'items peuvent être des solutions à l'élaboration de moyens d'évaluation équivalents. (Les banques d'items permettent d'accumuler un grand nombre de questions codifiées en fonction des objectifs qu'elles mesurent.)

LES MODÈLES MIXTES

Puisque nos systèmes d'éducation exigent habituellement une note différenciée afin de sélectionner les meilleurs étudiants et qu'il est par ailleurs difficile, dans le contexte d'un semestre, de maintenir un système d'évaluation formative et un système d'évaluation sommative, vu le temps requis pour la préparation et l'application des instruments de mesure, on essaye de faire des compromis entre les deux systèmes. Il faut cependant préserver la cohérence. Ceci n'est pas facile, car il s'agit de deux modèles d'évaluation visant des buts différents. Il n'y a sans doute pas de solution idéale: il s'agit de choisir celle qui convient le mieux au cours et au programme de formation.

Par exemple, la réussite à un test ou travail contribue pour tel pourcentage à la note, le test ou travail pouvant être repris sans pénalité jusqu'à ce que la note atteigne un standard prédéterminé ou que l'étudiant soit satisfait de ses résultats. D'une autre façon, on peut accorder une note (comme 60% ou 70%) pour la réussite des tests ou travaux formatifs et compléter cette note par celle qui est attribuée pour un examen ou travail supplémentaire permettant de différencier les étudiants. Ce ne sont que deux exemples parmi plusieurs possibles.

Conclusion

L'évaluation sommative et l'évaluation formative de l'apprentissage répondent à des fins différentes; elles servent d'appui à des décisions distinctes. Il ne s'agit donc pas de vanter les mérites d'un système d'évaluation par rapport à l'autre, mais de choisir un système d'évaluation répondant aux besoins identifiés.

Comme l'évaluation formative, surtout lorsqu'elle est structurée, n'est pas très répandue dans notre système d'éducation, il reste à faire plusieurs recherches et expérimentations en ce domaine. Ce texte ne peut donc pas être un tableau définitif des modèles d'évaluation.

Il est à noter que les modèles d'évaluation peuvent s'appliquer à d'autres objets que l'apprentissage de l'étudiant. On peut songer, dans le domaine de l'éducation, à l'évaluation de l'enseignement donné par un professeur[15] ou à l'évaluation d'un programme.

Par exemple, l'enseignement d'un professeur peut être évalué afin de décider s'il sera réengagé ou promu (évaluation sommative) ou encore afin de lui permettre d'améliorer son enseignement (évaluation formative). Ceci peut se faire en fonction de critères (évaluation critériée) ou en fonction des résultats de ses collègues (évaluation normative). L'évaluation de programme peut avoir pour but de choisir le ou les programmes qui pourront continuer à bénéficier de subsides, qui devront disparaître, etc. (évaluation sommative) ou encore déterminer les éléments du programme qui sont à réviser (évaluation formative). L'évaluation peut se faire en fonction de critères ou en fonction d'autres programmes.

Notes et références bibliographiques

1.Fontaine, France, «Il ne faut pas confondre l'autonotation et l'auto-évaluation» dans *Dossier sur l'évaluation*, pp. 29-32.

2.Fontaine, France, *Les objectifs d'apprentissage*. Montréal, Service pédagogique, Université de Montréal, 1980, 110 p..

3.Bloom, Benjamin S. et al., *Taxonomie des objectifs pédagogiques, Tome I, domaine cognitif*. Traduit de l'américain par: Lavallée, Marcel. Taxonomy of Educational Objectives, 1956. Montréal, Education nouvelle, 1969, 232 p..

4.Fontaine, France, «Évaluer? Mesurer?» dans *Dossier sur l'évaluation*, pp. 5-8.

5.Id., ibid.

6.Fontaine, France, «Il ne faut pas confondre l'autonotation et l'auto-évaluation» dans *Dossier sur l'évaluation*, pp. 29-32.

7.Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques*. Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.

8.Fontaine, France, *Les objectifs d'apprentissage*. Montréal, Service pédagogique, Université de Montréal, 1980, 110 p..

9.Felx-Larivière, Claire, Fontaine, France, «Assurer à tous la compétence» dans *Dossier sur l'évaluation*, pp. 11-15.

10.Id., ibid.

11.Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques*. Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.

12.Felx-Larivière, Claire, Fontaine, France, «Assurer à tous la compétence» dans *Dossier sur l'évaluation*. pp. 11-15.

13.Leclerc, Jean-Marc, *Dossier sur l'enseignement: systèmes, méthodes, techniques*. Montréal, Service pédagogique, Université de Montréal, 1977, pagination multiple.

14.Bernard, Huguette, Normand, Serge, «La banque d'objectifs et d'items: un exemple d'application» dans *Dossier sur l'évaluation*, pp. 123-130.

15.Fontaine, France, «Démarche pour évaluer l'enseignement» dans *Dossier sur l'évaluation*, pp. 51-71.

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Il ne faut pas confondre l'autonotation et l'auto-évaluation

France Fontaine

Que de fois, lors d'une discussion ou dans les faits, les gens ont confondu et confondent encore l'auto-évaluation et ce que nous appellerons l'autonotation. Ainsi, certains professeurs demandent à leurs étudiants de s'auto-évaluer, en attendant d'eux qu'ils s'attribuent une note, cette note étant plus ou moins sujette à révision avant d'être versée à leurs dossiers. Cette façon de procéder, qui eut plusieurs adeptes il y a quelques années et qui en compte encore un bon nombre, nous la désignerons sous le vocable «autonotation».

L'autonotation

Il existe plusieurs variantes de cette façon de procéder. Par exemple, la note attribuée par l'étudiant peut faire l'objet d'une discussion avec le professeur seul ou avec les étudiants du groupe. Il arrive souvent que cette note, qu'on espère la plus juste possible, soit donnée globalement par l'étudiant, c'est-à-dire sans qu'il se réfère à une liste de critères préétablis: une telle façon de procéder ne peut que rendre la note plus subjective. Lorsqu'il s'agit de groupes de travail, il peut y avoir attribution d'une même note à toute l'équipe ou encore l'on peut demander à chacun de s'attribuer une note et d'en accorder une à chaque membre de l'équipe en fonction du travail de chacun, la note de chaque étudiant étant par exemple la moyenne de ces notes pondérées ou pas.

On devine bien les dangers inhérents à l'autonotation et les situations aberrantes qu'elle peut entraîner ont maintes fois fait l'objet de discussions. Un des grands problèmes est dû au fait que les notes sont subjectives même si les étudiants sont honnêtes envers eux-mêmes. Certains ont tendance à se sous-estimer, tandis que d'autres se surestiment et il est bien difficile de départager ces deux catégories d'étudiants. La chose est d'autant plus grave lorsqu'il s'agit d'une classe nombreuse (ce qui se produit fréquemment à l'université) et que les mesures sur lesquelles le professeur peut se guider pour

apprécier les capacités des étudiants sont limitées quant au nombre. Dans ces cas, les étudiants ont pour ainsi dire le dernier mot, leurs notes pouvant difficilement être remises en question. D'ailleurs, peut-on vraiment dire qu'un étudiant s'accorde une note si le professeur a droit de véto. Les cas d'échec sont peu fréquents puisqu'il est rare qu'un étudiant ne s'accorde pas, à tout le moins, la note de passage et qu'il n'agisse pas ainsi à l'égard d'un coéquipier même si celui-ci a peu participé au travail: un étudiant refuse habituellement de donner une mauvaise note à un confrère qui a peu ou mal travaillé en se disant plutôt qu'on ne le reprendra plus à faire équipe avec cet individu.

Face à toutes ces difficultés qu'on attribue à tort à l'évaluation, faut-il s'étonner de l'attitude de plusieurs professeurs qui rejettent l'auto-évaluation pratiquée par les étudiants. Les dangers mentionnés précédemment sont tels, que nous comprenons sans peine la réaction de ces professeurs, mais si nous parvenons à distinguer l'auto-évaluation de l'autonotation, peut-être pourrons-nous démontrer qu'il peut être fort important d'amener l'étudiant à s'auto-évaluer adéquatement.

L'auto-évaluation

Lorsqu'il s'agit d'autonotation, le produit auquel on réfère est une note même si on mentionne qu'on espère aussi amener l'étudiant à être plus critique à l'égard de sa démarche. Certains professeurs ont même vu là l'occasion de se décharger de cette tâche administrative qu'est l'attribution des notes. Cependant, au lieu de viser comme résultat du processus la production d'une note qui (on sait le poids qu'elle a dans le dossier-étudiant) mène souvent à des situations aberrantes, on peut tendre à développer chez l'étudiant certaines capacités à s'évaluer en utilisant des critères définis préalablement. Les critères peuvent être communs à tous ou personnels s'il s'agit d'un enseignement personnalisé où chaque étudiant peut poursuivre des objectifs qui lui sont propres. Il est à noter que les critères communs ne restreignent en rien l'originalité, la créativité, puisque celles-ci peuvent servir de bases à des critères.

Il est plusieurs domaines où l'étudiant universitaire doit apprendre à garder un esprit critique face à ses capacités puisqu'il devra poursuivre seul sa formation, les connaissances et les techniques étant rapidement périmées. Par exemple, dans un cours de médecine, il est important que l'étudiant puisse porter un jugement sur l'acte médical qu'il vient de poser et que ce jugement soit aussi fiable que possible, puisque dès la fin de son cours, il n'y aura plus toujours quelqu'un pour lui dire d'améliorer ceci ou cela. De la même façon, l'étudiant en médecine dentaire devra être capable d'évaluer la prothèse qu'il a fabriquée puisqu'il aura à le faire dès qu'il pratiquera. Ce ne sont là que quelques domaines où l'auto-évaluation a son importance. On peut penser qu'il est plus facile de l'appliquer lorsqu'il y a production concrète, mais on doit aussi s'attendre, par exemple, à ce qu'un étudiant en psychologie clinique puisse évaluer adéquatement sa rencontre avec un patient, sa façon de mener l'entretien.

Cependant, on peut difficilement espérer que l'étudiant se montrera habile à s'auto-évaluer à la fin de son programme d'études si on ne l'a pas habitué à se jauger, si on ne l'a pas aidé à mieux s'auto-évaluer pendant ses études. Mais comment favoriser l'apprentissage de l'auto-évaluation?

Il est préférable, pour l'étudiant, de s'évaluer non pas globalement mais en fonction de critères précis: plus les critères sont précis, plus le jugement peut être valable. S'il est possible d'obtenir certaines mesures quant à ses capacités, cela lui facilitera d'autant plus la tâche. Son jugement devrait ensuite être comparé à un jugement-critère de façon à corriger, s'il y a lieu, son auto-évaluation. Les modalités pour guider l'apprentissage de l'étudiant en ce domaine peuvent être multiples. Cependant, il faut choisir avec circonspection puisque certaines modalités, à cause de leurs effets secondaires, entraînent des situations qui annihilent les résultats escomptés.

Par exemple, il serait possible d'imaginer qu'un étudiant en médecine juge chacun de ses actes médicaux en fonction des critères établis en fonction de sa profession, il pourrait même y avoir une grille élaborée à cet effet. Cette évaluation pourrait être comparée à l'évaluation faite par son ou ses superviseur(s). S'il n'y a pas concordance des deux évaluations, les points de dissemblance pourraient faire l'objet d'une discussion avec l'étudiant de sorte que ce feedback lui permettrait de comprendre le pourquoi de ces dissemblances et de s'auto-évaluer plus justement la fois suivante. Les superviseurs peuvent eux aussi tirer profit de ces discussions: ils peuvent, à l'occasion, repréciser la grille ou modifier une évaluation. Il est important de donner à l'étudiant la possibilité d'apprécier la valeur de son jugement, de l'aider à voir ses lacunes et à y remédier.

Il n'est absolument pas nécessaire que le processus mène à l'obtention d'une note contribuant à la note moyenne de l'étudiant. On a l'habitude de croire qu'on ne peut rien attendre des étudiants si on ne leur donne pas de notes. Il est cependant possible, surtout s'il s'agit d'un enseignement par objectifs liés à une évaluation formative, d'exiger des étudiants qu'ils soient capables de s'auto-évaluer avant de considérer que leur cours est réussi. Il s'agit alors d'un objectif obligatoire à atteindre: on a mis en oeuvre des activités pédagogiques pour y arriver, les moyens d'évaluer si l'objectif est atteint ou pas et les moyens de fournir un feedback favorisant l'amélioration.

Conclusion

L'autonotation et l'auto-évaluation sont des notions qui peuvent être très distinctes. Si elles étaient mieux comprises, il y aurait peut-être un plus grand nombre de programmes d'études accordant une importance dans les faits à l'auto-évaluation au cours de la formation de leurs étudiants.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Bulletin*, Montréal, Service pédagogique, Université de Montréal, No 4, Oct. 1975.

Section 2

Des processus d'élaboration d'instrument d'évaluation

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

L'utilisation des objectifs pour évaluer l'apprentissage lors d'un stage

Huguette Bernard
France Fontaine

Lorsque des professeurs concernés par les stages ou d'autres formes d'enseignement pratique se trouvent réunis, on les entend souvent mentionner, à un moment ou à un autre, que l'évaluation de l'apprentissage des étudiants est un de leurs problèmes. «Évaluer» leur semble d'autant plus difficile qu'il leur est impossible de se rabattre sur un «examen papier-crayon» (qu'on croit, à tort, facile à bien construire) pour mesurer les apprentissages pratiques.

Certains professeurs essayent donc d'améliorer l'évaluation des acquis chez les étudiants en modifiant, d'année en année, leurs moyens d'évaluer. Leurs démarches sont très souvent circulaires, on planifie plus ou moins les transformations: on modifie une grille pour revenir à l'ancienne grille quelques années plus tard, on rédige des objectifs sans trop savoir ce qu'on en fera ensuite et ainsi de suite.

C'est pour aider ces gens intéressés à élaborer *un des moyens d'évaluer l'apprentissage*, c'est-à-dire *une grille d'évaluation* qui correspond aux *objectifs de stage*, que nous présentons un processus de production d'une telle grille.

Il existe plusieurs types de grilles d'évaluation, mais nous nous intéresserons uniquement à celles qui font le lien entre les objectifs du stage et ce qu'on évalue. Ainsi, la première partie de ce texte vise à exposer *une démarche favorisant la rédaction et l'organisation des objectifs de stage*. La seconde partie portera sur *le processus d'élaboration d'une grille d'évaluation de l'apprentissage*.

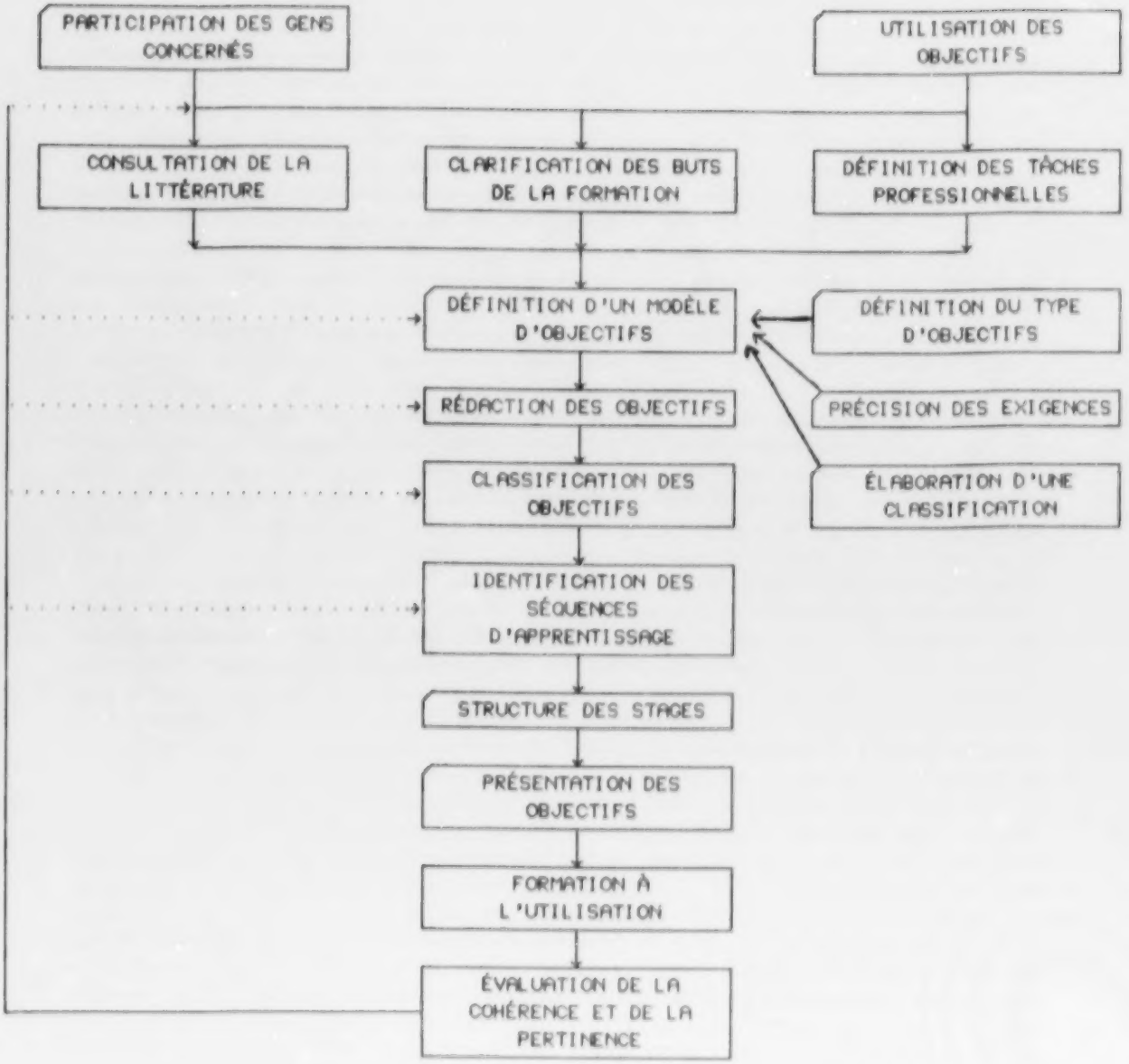
Cette présentation des étapes de rédaction des objectifs et de production d'une grille d'évaluation d'un stage est inspirée de notre expérience de conseillers pédagogiques auprès de professeurs qui se sont intéressés aux problèmes des stages. C'est la réflexion à partir de ces expériences qui nous a permis d'en clarifier le processus.

LA RÉDACTION ET L'ORGANISATION DES OBJECTIFS DE STAGE

Il serait irréaliste de penser que l'on puisse proposer une manière de rédiger des objectifs de stage qui serait la même pour tous. Cependant, on peut tenter de faire ressortir certains points favorisant une rédaction valable des objectifs.

Le tableau qui suit présente l'organigramme des étapes à suivre en vue de rédiger et d'organiser des objectifs de stage. On y retrouve des décisions à prendre et des activités de production. Cette partie du texte sera entièrement consacrée à expliciter ce tableau. Notons cependant qu'il y a plusieurs étapes préalables à la rédaction des objectifs, étapes qu'on omet trop souvent d'ailleurs. Rédiger des objectifs ne suffit pas non plus: si on veut que ceux-ci soient utilisés efficacement, il faut aussi les organiser.

DÉMARCHE VISANT À RÉDIGER ET À ORGANISER DES OBJECTIFS DE STAGE*



* Les rectangles tronqués indiquent qu'il y a une décision à prendre.

La participation des gens concernés

Il est important de souligner qu'on doit s'assurer la participation des diverses personnes concernées: professeurs, étudiants, cliniciens ou moniteurs, administrateurs...

Il serait dangereux, par exemple, qu'un groupe restreint de professeurs rédige les objectifs sans consulter les autres professeurs concernés. Il est difficile pour les personnes qui n'ont pas participé à la rédaction des objectifs de se rallier à ces objectifs et, par la suite, d'en tenir compte lors d'un stage. Dans le cas où on désirerait que les étudiants choisissent des objectifs personnels en «pigeant» dans une banque d'objectifs, il serait inconséquent de ne pas les faire participer à l'élaboration de cette banque.

Evidemment, la participation peut se jouer à divers niveaux: certaines personnes peuvent être plus engagées que d'autres, mais il est avantageux que l'information circule rapidement et qu'il y ait place pour la critique et la rectification.

Dans certains cas, on doit former un comité composé de représentants de chacune des instances concernées et fixer le nombre de représentants de façon à équilibrer le comité de rédaction. Ceux qui sont choisis doivent être très motivés car, en plus de la tâche qui les attend, ils ont à informer et à susciter l'intérêt de ceux qu'ils représentent.

L'utilisation des objectifs

Nous avons, à maintes reprises, observé que des objectifs rédigés soit pour des cours, soit pour des stages, n'ont pas eu d'impact sur ces cours ou ces stages. Or, il faut bien se rappeler que la rédaction d'objectifs précis influence véritablement l'enseignement dans la mesure où il y a interaction entre ces objectifs et les autres facettes de l'acte éducatif (c'est-à-dire les méthodes d'enseignement, les média d'enseignement, les techniques d'évaluation, etc.)

Il est important qu'il y ait un but à la rédaction des objectifs, que l'on se propose d'utiliser ceux-ci de façon concrète. Il pourrait être dangereux d'entreprendre la rédaction des objectifs sans avoir prévu ce qu'on en fera; ils pourraient rester inexploités, comme ces objectifs que l'on conçoit pour suivre une certaine mode dans le domaine de l'éducation ou pour répondre à des exigences administratives. L'utilisation qu'on veut en faire peut aussi être un facteur déterminant lorsqu'il faut choisir le modèle des objectifs à rédiger.

Même si dans ce texte on a déjà choisi *une utilisation* des objectifs qui est *l'élaboration d'une grille d'évaluation* qui correspond à ces objectifs, il serait opportun de souligner quelques-unes des autres utilisations possibles.

Pour assurer une continuité dans l'apprentissage

En ce qui concerne chacun des objectifs ou des sous-groupes d'objectifs relatifs à un stage, on peut préciser quels en sont les prérequis, c'est-à-dire les connaissances et aptitudes préalables. Le professeur ou l'étudiant lui-même (si le professeur lui fournit les instruments) peut ensuite vérifier si les prérequis sont maîtrisés ou pas. L'étudiant peut combler ses lacunes, s'il y a lieu; il peut le faire plus facilement s'il y a des prescriptions d'activités concernant chacun des prérequis non maîtrisés.

Pour assurer la séquence des stages

La rédaction d'objectifs précis suppose qu'on les étudie ensuite afin d'y faire apparaître, s'il y a lieu, des séquences d'apprentissage, c'est-à-dire des objectifs qui doivent être maîtrisés dans un certain ordre. Les objectifs facilitent donc la structuration des stages: par exemple, l'objet du premier stage, ce qui fait partie du second, quels sont les apprentissages du début du stage et le reste.

Pour déterminer les activités d'apprentissage

Plusieurs professeurs se servent des objectifs pour choisir leurs méthodes d'enseignement, les activités d'apprentissage du stage. Par exemple, lorsqu'on identifie bien les objectifs se rapportant à la communication, il est plus facile de penser à des activités précises favorisant ces apprentissages.

Pour rendre la compétence accessible à tous

Lorsqu'on considère que la compétence n'est pas le privilège d'une minorité mais doit être accessible à tous, on cherche à fournir à tous les étudiants, à l'aide de moyens variés, la possibilité d'acquérir la compétence. Il faut donc, dans un premier temps, préciser les objectifs, c'est-à-dire les apprentissages visés et les standards indiquant la compétence. On peut ensuite établir les moyens pour atteindre ces objectifs et pour les contrôler[1].

La consultation de la littérature

Il doit y avoir vérification du matériel déjà existant. Il peut s'agir de listes d'objectifs déjà existantes dans le département ou dans un département donnant une formation analogue dans une autre université. Les objectifs provenant d'une autre discipline peuvent aussi servir de guide sur le plan du modèle utilisé pour la rédaction des objectifs ou pour l'utilisation qu'on en fait.

La définition des tâches professionnelles

Lorsqu'on structure l'enseignement des stages, on s'inspire de l'acte professionnel pour rédiger les objectifs. On peut avoir recours à une analyse de besoins pour déterminer ce qu'on attend du professionnel de l'avenir: on s'inspire alors non pas uniquement des tâches actuelles mais des tâches futures de ce professionnel.

La clarification des buts de la formation

Il serait dangereux de vouloir rédiger des objectifs précis sans d'abord s'interroger sur les buts de la formation. Cette façon de procéder permet d'éviter les objectifs non pertinents grâce à la confrontation constante de ceux-ci avec les buts de la formation. On écarte ainsi le danger de rejeter une catégorie d'objectifs parce qu'on ne parvient pas à les préciser de façon satisfaisante, ou bien de retenir des objectifs simplistes mais faciles à rédiger.

Cette réflexion complète bien la définition des tâches professionnelles puisqu'elle indique de quelle façon on veut tenir compte des tâches actuelles et/ou futures du professionnel dans la formation universitaire.

La définition d'un modèle d'objectifs

Il est important de préciser la forme que prendront les objectifs. Sinon, on risquerait de se retrouver avec une liste presque inutilisable d'objectifs hétéroclites.

Les contingences de la profession et les buts de la formation déterminent en grande partie le modèle choisi. Par exemple, pour ce qui est d'une formation clinique donnée, si on considère que chacun des patients doit être considéré dans son ensemble et non pas en fonction de sa pathologie, on rédigera des objectifs parmi lesquels ne figurera pas la nomenclature des diverses pathologies à traiter. Si, par contre, il est essentiel que l'étudiant soit capable de traiter tel ou tel cas pathologique, on devra soumettre l'étudiant à un contrôle qui tienne compte de cet aspect de sa formation et on devra aussi rédiger les objectifs de façon à rendre ce contrôle possible.

Lorsqu'on choisit un modèle de rédaction des objectifs, on doit aussi penser à l'utilisation des objectifs. Si le fait d'écrire les objectifs de telle façon plutôt que telle autre génère plusieurs centaines d'objectifs pour un stage, on devra songer à choisir un autre modèle afin de pouvoir les utiliser.

Lorsqu'on définit le modèle désiré lors de la rédaction des objectifs, il faut, par exemple, décider des types d'objectifs, décider si tous les objectifs seront également poursuivis dans la formation, élaborer une classification des objectifs.

La définition du type d'objectifs

Lorsqu'on veut rédiger des objectifs précis de stage, c'est-à-dire des *objectifs spécifiques*[2], on écrit des énoncés qui décrivent ce vers quoi tend l'apprentissage, en termes de comportement observable. Ces objectifs, dits spécifiques, doivent répondre à des règles précises: par exemple, ils sont exprimés en fonction de l'étudiant, ils n'acceptent qu'une seule interprétation. On conseille de rédiger ces objectifs spécifiques en utilisant des énoncés moins précis qu'on nomme *objectifs généraux* et qui décrivent, en termes plus vagues, ce vers quoi tend l'apprentissage. Ils constituent une première ébauche de ce qu'on attend de l'étudiant.

Un objectif général peut générer plusieurs objectifs spécifiques. Autrement dit, un objectif spécifique est une interprétation de l'objectif général qui le sous-tend. Voici un exemple d'objectif général et de quelques-uns des objectifs spécifiques qui peuvent en découler.

Objectif général :

- L'étudiant en géologie devra connaître les divers minéraux.

Objectifs spécifiques :

- L'étudiant en géologie devra, pour chaque nom de minerai donné, indiquer les caractéristiques premières de chacun dans une brève description.
- L'étudiant en géologie devra pouvoir indiquer le degré de dureté de chaque minerai donné.
- L'étudiant en géologie devra, en laboratoire, identifier un minerai à partir de trois spécimens différents de ce minerai.
- L'étudiant en géologie devra, lorsque mis en présence d'un minerai sur le terrain, identifier sans faute celui-ci.

On voit que plusieurs objectifs spécifiques peuvent être déduits d'un même objectif général. Cependant, tous ne sont pas nécessairement pertinents. Ainsi, le dernier objectif spécifique peut être le seul objectif qu'on veuille retenir lors d'un stage (en géologie) sur le terrain. Les autres objectifs spécifiques mentionnés, ou d'autres objectifs spécifiques rédigés à partir du même objectif général, peuvent être considérés comme des étapes pour atteindre cet objectif spécifique de stage, qu'on peut aussi appeler *objectif terminal*.

Chacun des autres objectifs peut être appelé un *objectif intermédiaire* car il indique, dans une séquence d'apprentissage, une étape que le sujet doit franchir avant d'atteindre l'objectif terminal. L'objectif terminal étant cette capacité durable et intrinsèquement utile que le sujet doit avoir acquise à la fin d'un apprentissage déterminé.

Il faut donc décider, lorsqu'on rédige les objectifs spécifiques d'un stage, sous quelle forme on les présentera. Indiquera-t-on l'objectif général suivi des objectifs spécifiques qui en découlent? Indiquera-t-on plutôt les objectifs spécifiques dans un ordre d'apprentissage, c'est-à-dire l'objectif terminal et les objectifs intermédiaires qui y correspondent?

La précision des exigences

Il faut décider, avant ou après la rédaction des objectifs, si tous les objectifs ont une même importance dans la formation: sont-ils tous obligatoires ou prioritaires ou en existe-t-il que l'on considère comme étant optionnels ou secondaires? Certains objectifs doivent-ils faire partie de la formation de tout étudiant tandis que d'autres peuvent être définis ou choisis en fonction des besoins personnels de chacun des étudiants?

L'élaboration d'une classification des objectifs

Il est important de classer les objectifs suivant les comportements qui leur sont rattachés. Lorsqu'on omet cette étape, on se retrouve avec une simple liste d'objectifs sans aucun regroupement. Une telle liste est pratiquement inutilisable, car il est difficile d'y repérer rapidement un objectif.

Il faut donc choisir ou élaborer une classification des objectifs. On peut utiliser les taxonomies plus connues du domaine cognitif, affectif et psychomoteur[3]. On peut aussi organiser les objectifs selon les tâches accomplies par le professionnel. On peut, dans le domaine médical par exemple, retrouver des catégories portant sur des capacités de faire l'histoire de cas, le diagnostic, le traitement et le reste.

La rédaction et la classification des objectifs

Une fois que la décision quant au choix du modèle des objectifs a été prise, il ne reste plus qu'à se mettre au travail, c'est-à-dire à rédiger les objectifs et à les classer. On peut veiller à ne pas oublier d'objectifs en les confrontant avec la classification choisie, les buts de la formation et les tâches professionnelles.

La rédaction des objectifs est un point central du processus, mais ce n'est cependant pas le point terminal: il reste plusieurs étapes à accomplir si on veut que la rédaction des objectifs soit vraiment rentable.

L'identification des séquences d'apprentissage

Après avoir rédigé et classé les objectifs de stage, il faut voir, comme nous le mentionnions précédemment, s'il est possible de les organiser en séquences d'apprentissage. Ceci permet de structurer les stages de façon à favoriser un meilleur apprentissage. Il est plus facile, par exemple, pour l'étudiant et/ou pour le moniteur (ou clinicien) de choisir les activités d'apprentissage les plus pertinentes au début d'un stage. Cependant, ceci n'est pas toujours facile à réaliser: les contingences de certains endroits où se déroulent les stages ne permettent pas toujours d'en tenir compte.

Il est nécessaire d'identifier les séquences d'apprentissage non pas uniquement à l'intérieur des stages mais dans tout le programme: on indique les connaissances théoriques préalables. Par exemple, un apprentissage ayant comme prérequis la connaissance de telle notion requérant deux années de formation ne peut faire l'objet d'un stage dès la première année.

La structuration des stages

Une fois les séquences d'apprentissage identifiées, il est plus facile de déterminer le cadre des stages, c'est-à-dire les regroupements d'apprentissages qui feront l'objet d'un stage, le temps requis pour chacun des stages, le nombre de crédits par stage et le reste.

Ici, bien sûr, nous avons adopté une optique bien particulière où il est possible de réorganiser l'ensemble des activités de stage relatives à un programme. Plusieurs n'ont malheureusement pas cette possibilité de pouvoir corriger certaines incohérences et ne peuvent que modifier un stage bien précis.

La présentation des objectifs

C'est la phase finale de la production. On définit sous quelle forme les objectifs seront présentés. L'utilisation qu'on fera des objectifs conditionnera sans doute leur présentation: guide de stage, programme, grille d'évaluation... Le format est très important lorsque l'étudiant et le moniteur s'en servent fréquemment; la présentation visuelle l'est tout autant.

La formation à l'utilisation des objectifs

Une fois les objectifs rédigés, il faut songer à les utiliser efficacement. Des instructions et des sessions de formation de plus ou moins courte durée doivent être prévues à cet effet, surtout dans le cas de personnes qui n'ont pas ou peu participé à l'élaboration des objectifs.

La formation doit viser les professeurs, les moniteurs ou cliniciens et les étudiants. Il devient important d'expliquer, à l'étudiant, la place qu'occupent les objectifs dans son processus d'apprentissage.

L'évaluation de la cohérence et de la pertinence des objectifs

S'il y a des vérifications importantes à faire après chacune des étapes, c'est l'expérimentation des objectifs de stage qui est sans doute la grande épreuve.

On vérifie si les cours et les divers stages s'enchaînent bien, si les apprentissages sont cohérents et pertinents, compte tenu de la formation visée. Il y a toujours des modifications à faire. Parfois, elles sont mineures: quelques objectifs doivent être clarifiés par exemple. Parfois, elles sont majeures: des modifications quant à l'utilisation des objectifs ou encore des changements dans les buts de la formation pratique dus à des transformations rapides dans la profession. Sur le plan de l'utilisation des objectifs, notons qu'il est parfois possible de procéder à des changements sans investir nécessairement beaucoup de temps.

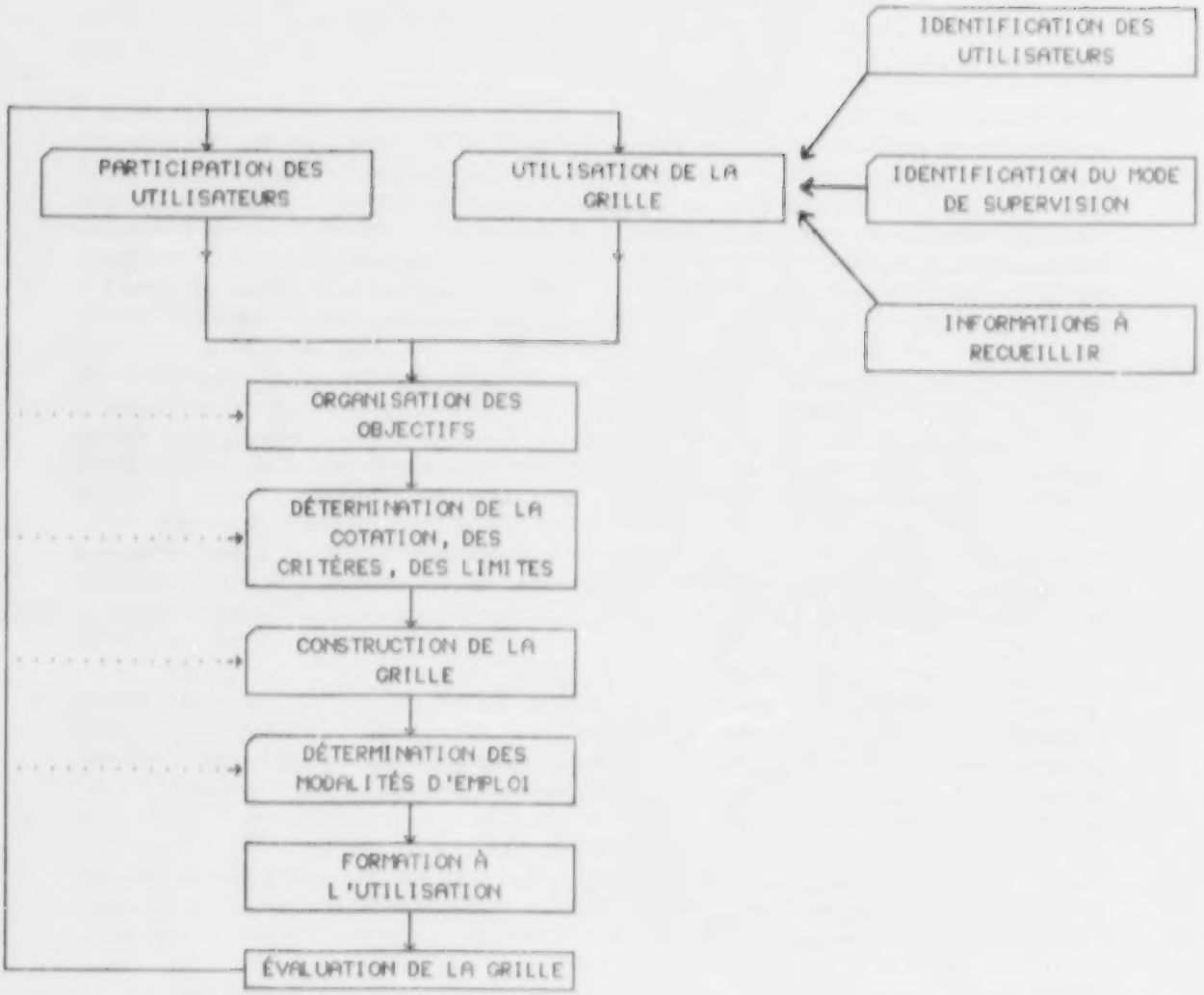
L'UTILISATION DES OBJECTIFS DANS UNE GRILLE D'ÉVALUATION DE L'APPRENTISSAGE

Nous nous attarderons ici à *une des utilisations des objectifs de stage*. Plus précisement, il sera question de l'intégration des objectifs à une grille d'évaluation de l'apprentissage. L'évaluation est structurée, organisée en fonction de ces objectifs; nous laissons de côté dans ce texte l'évaluation non structurée comme le feedback spontané qu'un moniteur donne à un étudiant et qui peut concerner d'autres aspects que ceux traités dans les objectifs.

De la même manière dont nous avons procédé pour décrire les étapes concernant la rédaction et l'organisation des objectifs de stage, nous indiquerons, en ce qui a trait aux stages, certains éléments facilitant l'élaboration d'une grille d'évaluation de l'apprentissage.

Le tableau suivant résume les étapes relatives à la production d'une grille pour évaluer les apprentissages d'un stagiaire. Certaines de ces étapes constituent des décisions à prendre, d'autres des activités de production. Cette seconde partie du texte explique ce tableau.

DÉMARCHE POUR ÉLABORER UNE GRILLE D'ÉVALUATION DE L'APPRENTISSAGE*


IDENTIFICATION DES UTILISATEURS
PARTICIPATION DES UTILISATEURS
UTILISATION DE LA GRILLE
IDENTIFICATION DU MODE DE SUPERVISION
INFORMATIONS À RECUEILLIR
ORGANISATION DES OBJECTIFS
DÉTERMINATION DE LA COTATION, DES CRITÈRES, DES LIMITES
CONSTRUCTION DE LA GRILLE
DÉTERMINATION DES MODALITÉS D'EMPLOI
FORMATION À L'UTILISATION
ÉVALUATION DE LA GRILLE

* Les rectangles tronqués indiquent qu'il y a une décision à prendre.

La participation des utilisateurs de la grille

Il est important de sensibiliser, de consulter, de faire participer les utilisateurs (étudiants, professeurs, moniteurs, etc.) d'une grille d'évaluation de l'apprentissage. On peut ainsi s'assurer que la grille correspond bien à leurs besoins. De plus, une telle façon de procéder facilite l'emploi ultérieur de la grille.

L'utilisation de la grille

Lorsqu'on a choisi la construction d'une grille d'évaluation de l'apprentissage comme but de la rédaction des objectifs de stage, on a sans doute précisé de *quel type d'évaluation* il s'agissait (une évaluation formative ou sommative, un cadre de référence normatif ou critérié)[4]. Sinon, il est important de clarifier le pourquoi de l'évaluation: Est-ce pour rendre l'étudiant capable de s'autocritiquer? Est-ce pour assurer la certification des étudiants? Est-ce pour donner du feedback à l'étudiant en fonction des objectifs? Est-ce pour améliorer la qualité du stage en identifiant les objectifs rarement atteints, les stratégies d'enseignement à réviser? Est-ce pour mieux administrer les stages en identifiant les objectifs qui ne peuvent être atteints à tel endroit où se déroule un stage? Est-ce une combinaison de ces buts qu'on poursuit en voulant évaluer? Notons cependant que certains buts sont parfois contradictoires: certaines combinaisons sont irréalisables.

Pour ce qui est de la grille, les réponses données aux questions suivantes jettent aussi un peu de lumière sur son utilisation: qui utilisera la grille? quel sera le mode de supervision? quelles informations veut-on recueillir?

L'identification des utilisateurs de la grille

Parmi les utilisateurs éventuels d'une grille d'évaluation, on trouve l'étudiant lui-même, des étudiants ayant une même formation ou une formation plus avancée, un moniteur ou un professeur. La grille peut être utilisée par une ou plusieurs de ces catégories de personnes.

L'identification du mode de supervision

Il peut être utile de se demander si le mode de supervision utilisé ou souhaité est direct ou indirect. (La supervision directe offre bien des avantages, mais peut parfois biaiser la relation client/étudiant dans le cas où on n'utilise pas le miroir unidirectionnel ou le magnétoscope; la supervision indirecte basée uniquement sur la discussion avec l'étudiant ne permet pas de donner du feedback pour toutes les catégories d'objectifs.)

On peut aussi décider si l'on veut faire de la supervision individuelle ou collective. En supervision collective, on peut trouver un ou plusieurs moniteurs (cliniciens, professeurs) avec quelques étudiants: chacun des étudiants reçoit alors du feedback de chacune des personnes du groupe et non pas uniquement de son moniteur.

Les informations à recueillir

En précisant les informations qu'on veut recueillir et à qui ces informations seront transmises, on clarifie à quelles fins on veut utiliser la grille d'évaluation. Désire-t-on savoir si l'étudiant atteint ou pas un objectif donné, si tel objectif peut être atteint ou pas dans chacun des endroits où il effectue un stage, si tel objectif est rarement atteint par les étudiants et ainsi de suite. Les réponses à ces questions permettent de clarifier ce qu'on attend de la grille d'évaluation.

L'organisation des objectifs

Si cela n'est pas déjà fait, comme nous l'avons proposé dans la partie du texte portant sur la rédaction des objectifs, il faut organiser une présentation des objectifs qui favorise le repérage rapide d'un de ceux-ci. Par exemple, on peut présenter des regroupements d'objectifs en fonction des diverses étapes de la démarche du professionnel; il est alors plus facile de repérer le regroupement, puis l'objectif précis qu'on veut évaluer dans ce bloc d'objectifs.

La détermination de la cotation, des critères, des limites

Même si on a une liste d'objectifs organisés, cela ne suffit pas à en faire une grille d'évaluation. Il faut savoir comment on cotera la performance de l'étudiant. Ainsi, selon qu'on favorise une évaluation sommative ou formative, il est plausible de coter l'objectif différemment (exemple: une cote de 1 à 5 pour une évaluation sommative et succès-échec avec prescription d'activités d'apprentissage pour une évaluation formative). Si on veut suivre l'évaluation de l'apprentissage d'un étudiant dans le temps, on devra coter en conséquence (exemple: pour un objectif, plusieurs cotes distinctes dans le temps plutôt qu'une seule).

Lorsqu'il s'agit d'une évaluation menée à des fins de certification, c'est-à-dire qui permet de juger si un étudiant a réussi ou pas son stage, on doit décider comment on fera le lien entre les cotes accordées à chacun des objectifs et la cote à accorder pour l'ensemble du stage. Les façons de procéder sont des plus variées pour répondre à divers buts, pour tenir compte de contraintes différentes. Par exemple, un stage pourra être réussi si tous les objectifs obligatoires sont atteints. Dans le cas où on veut différencier les étudiants, certains additionnent les cotes accordées à chacun des objectifs après avoir pondéré ou pas l'importance de chacun des objectifs.

On peut désirer que l'objectif comporte un critère comme un taux de réussite ou un nombre de fois où l'objectif doit être atteint. Il faut alors déterminer ces critères pour chacun des objectifs en tenant compte de la réalité, indiquer ceux-ci dans la grille qui devra être structurée de façon à pouvoir faciliter la comptabilisation.

Dans certains cas, on veut exiger qu'à tel moment donné un objectif ou un ensemble d'objectifs soit(ent) atteint(s). Il faut donc préciser ces limites de temps, les indiquer sur la grille ou sur un document complémentaire.

La construction de la grille d'évaluation

Il s'agit de l'étape de mise en forme de la grille d'évaluation en tenant compte de toutes les décisions précédentes, en vérifiant s'il y a des contradictions entre celles-ci, en rectifiant s'il y a lieu. On doit aussi penser à une présentation visuelle et à un format pratique.

La détermination des modalités d'emploi

On peut vouloir préciser ceci : quand remplir la grille, à quelle fréquence on a un rendement maximum et le reste. L'expérimentation de la grille peut donner des indications à ce sujet.

Dans le cas de l'évaluation sommative, on peut évaluer de façon à tenir compte des difficultés de l'étudiant au début de son stage : pour lui laisser le temps d'apprendre, on ne cumule pas dès le début ses succès et ses échecs par exemple.

La formation à l'utilisation de la grille

On peut songer à donner une formation plus ou moins complète aux utilisateurs de la grille. On doit à tout le moins s'assurer que tous comprennent ses buts et la façon de l'utiliser. Lorsqu'on compte des moniteurs (ou cliniciens) parmi les utilisateurs de la grille, on peut songer à leur apprendre à donner du feedback et à discriminer les comportements étudiants décrits dans les objectifs de la grille.

L'évaluation de la grille

Durant l'expérimentation de la grille, on accumulera des données permettant d'apporter des modifications : changer la présentation de la grille, le mode de cotation et le reste. C'est le fait d'un processus dynamique.

Conclusion

Ce texte n'a pas la prétention de faciliter la rédaction des objectifs de stage et l'élaboration d'une grille d'évaluation. Tout au plus vise-t-il à exposer un processus pour y arriver, en espérant qu'il y ait ainsi plus de cohérence dans les étapes. Pour illustrer ce texte par des exemples et du même coup le compléter, vous pourrez vous référer aux descriptions des expériences qui ont été faites en ce domaine par les départements de gynécologie-obstétrique[5] et d'ergothérapie[6] de l'Ecole de Réadaptation tous deux de la Faculté de Médecine de l'Université de Montréal.

Notes et références bibliographiques

1 Felx-Larivière, Claire, Fontaine, France, «Assurer à tous la compétence» dans *Dossier sur l'évaluation*, pp. 11-15

2 Vous trouverez l'explication des termes objectif général, objectif spécifique, objectif terminal et objectif intermédiaire dans le texte suivant : Fontaine, France, *Les objectifs d'apprentissage*. Montréal, Service pédagogique, Université de Montréal, 1980, 110 p.

3 Bloom, Benjamin S. et al., *Taxonomie des objectifs pédagogiques, Tome 1, domaine cognitif*. Traduit de l'américain par : Lavallée, Marcel. Taxonomy of Educational Objectives, 1956. Montréal, Education nouvelle, 1969, 232 p.

Krathwohl, David R., Bloom, Benjamin S., Masia, Bertram B., *Taxonomie des objectifs pédagogiques, Tome 2, domaine affectif*. Traduit de l'américain par : Lavallée, Marcel. Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook II : Affective Domain, 1964. Montréal, Education nouvelle, 1970, 231 p.

Harrow, Anita ., *Taxonomie des objectifs pédagogiques, Tome 3, domaine psychomoteur*. Traduit de l'américain par : Lavallée, Marcel. A Taxonomy of the Psychomotor Domain. A Guide for Developing Behavioral Objectives, 1972, Montréal, Les Presses de l'Université du Québec, 1977, 125 p.

De Landsheere, Gilbert, De Landsheere, Viviane, *Définir les objectifs de l'éducation*. Paris, Presses universitaires de France, 1977, 283 p.

4 Fontaine, France, «Évaluer? Mesurer?» dans *Dossier sur l'évaluation*, pp. 5-9.

5 Bernard, Huguette, Normand, Serge, «La banque d'objectifs et d'items : un exemple d'application» dans *Dossier sur l'évaluation*, pp. 123-130.

6 Forland, Francine, Hachey, Raymonde, «L'évaluation de la formation clinique en ergothérapie» dans *Dossier sur l'évaluation*, pp. 131-139.

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Démarche pour évaluer l'enseignement

France Fontaine

Lorsqu'on désire évaluer l'enseignement, il faut y consacrer le temps nécessaire. Il est inutile de brûler des étapes. Par exemple, on risquerait de recueillir des informations plus ou moins utiles ou de provoquer des conflits de travail en choisissant un instrument sans établir préalablement une politique d'évaluation. Il est donc très important de décider d'abord d'une *politique d'évaluation* et de déterminer ensuite les procédures du programme d'évaluation.

Il faut, en premier lieu, s'assurer la coopération des gens concernés, c'est-à-dire les professeurs, les étudiants et les autorités. L'expérience relative à plusieurs institutions nous l'a démontré: si l'ambiance du milieu de travail est malsaine, si on se sert de l'évaluation pour donner une apparence scientifique à des décisions prises préalablement, même si des instruments technologiquement bien conçus sont utilisés, il faut s'attendre à en tirer peu de profit.

En plus d'obtenir la *coopération du milieu*, il serait souhaitable d'assurer la circulation de *l'information* dans l'institution durant la mise sur pied d'un programme d'évaluation de l'enseignement. Ceci permet de susciter l'intérêt des gens et de les rendre responsables de la bonne marche du programme. Bref, c'est un premier gage d'efficacité.

POLITIQUE D'ÉVALUATION DE L'ENSEIGNEMENT

Voici maintenant les principales questions à considérer lors de l'élaboration d'une *politique d'évaluation de l'enseignement*

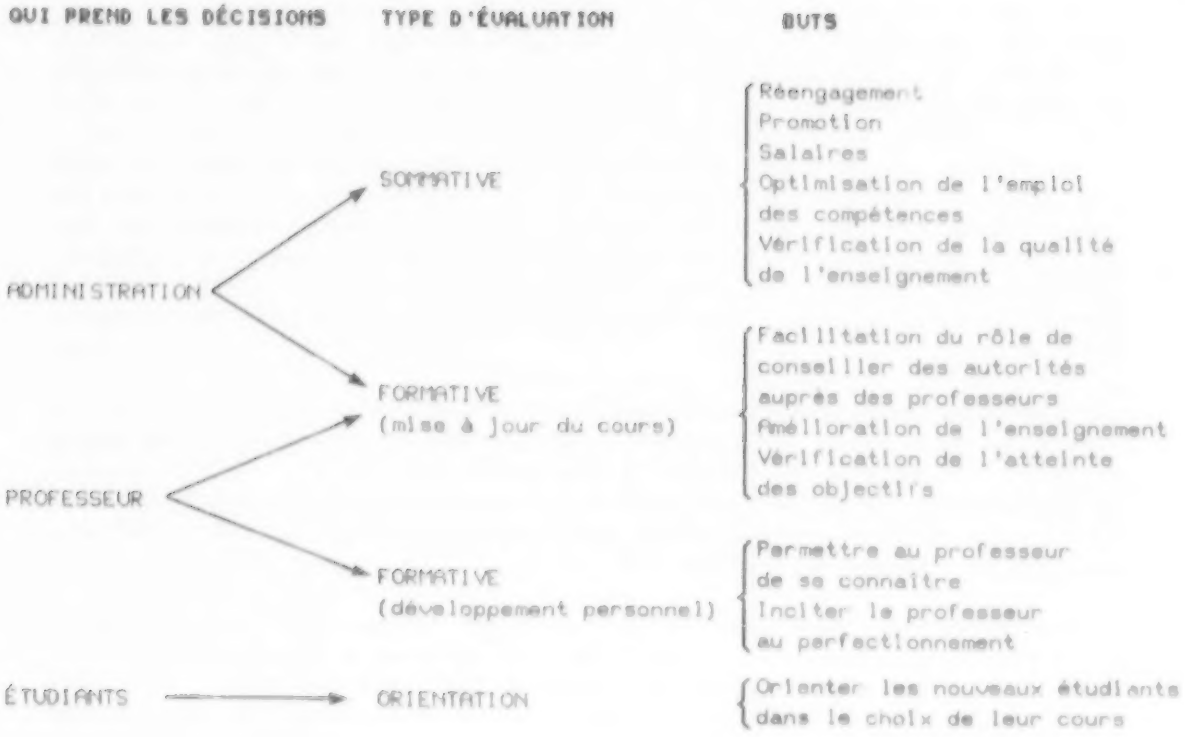
Pourquoi évaluer l'enseignement?

La première chose à faire est de préciser le pourquoi de l'évaluation, c'est-à-dire: *dans quel but veut-on évaluer?* Cette décision détermine en grande partie les choix à faire quant à l'objet de l'évaluation, aux instruments d'évaluation, à l'utilisation que l'on fera des résultats et ainsi de suite. C'est la décision préalable à toute démarche. S'il y a eu, dans plusieurs institutions, des réticences concernant l'évaluation de l'enseignement, c'est souvent dû au fait que les buts de l'évaluation étaient obscurs. De plus, s'assurer que les buts sont acceptés par les personnes concernées est un présage de succès.

Enumérons ici quelques buts possibles d'une évaluation de l'enseignement:

- Accélérer ou retarder l'obtention d'une promotion.
- Déterminer les augmentations de salaire au mérite.
- Déceler les cas marginaux qui ne sont pas à leur place dans l'enseignement, c'est-à-dire diminuer la part d'arbitraire dans le réengagement ou le non-réengagement d'un professeur.
- Vérifier si un professeur satisfait aux attentes préalablement établies par l'institution.
- Optimiser l'emploi des compétences en servant de support aux décisions relatives à l'affectation des professeurs à des tâches qui leur sont plus appropriées.
- Permettre au directeur ou au doyen de mieux conseiller les professeurs.
- Améliorer la qualité de l'enseignement. Provoquer un dialogue sur l'importance ou la pertinence de la matière et du programme, sur l'efficacité des moyens didactiques employés et le reste. Déceler les écarts entre les attentes des étudiants, les intentions du professeur et les besoins de la discipline afin de réduire ces écarts.
- Vérifier si les objectifs d'enseignement d'un cours ou d'un programme sont atteints.
- Identifier les problèmes relatifs à l'ajustement des unités de cours par rapport à l'ensemble d'un programme d'étude.
- Orienter les nouveaux étudiants quant au choix de certains cours selon l'appréciation qu'en ont faite les étudiants qui les ont suivis précédemment.

Le tableau qui suit permet de relier les *buts poursuivis* par une évaluation de l'enseignement au *type d'évaluation*[1] et à *la personne qui prend les décisions*.



Il se peut qu'on veuille poursuivre plusieurs buts à la fois en procédant à une évaluation de l'enseignement. Il faudra alors concevoir des programmes d'évaluation distincts, à moins que les buts poursuivis ne soient connexes. Il n'est pas rare de voir des programmes d'évaluation rencontrer des difficultés énormes parce qu'ils poursuivent des buts contradictoires.

Qui participe à l'élaboration du programme d'évaluation?

Dépendant du contexte dans lequel s'inscrit l'institution d'enseignement, de l'initiateur du projet et le reste, cette question peut se poser avant ou après la précédente (Pourquoi évaluer l'enseignement?). Il faut se demander qui élaborera la politique d'évaluation et qui participera à la mise en route et à la révision du programme d'évaluation. Par exemple, on peut retrouver des représentants des corps suivants:

- professeurs,
- étudiants,
- autorités,
- syndicats,
- professionnels de l'évaluation.

Si un comité est formé, il faut fixer le nombre de représentants de façon à équilibrer le comité. Il faut aussi songer à choisir un leader, une façon d'effectuer la consultation auprès des participants, un mécanisme de décision et ainsi de suite.

Il est avantageux de faire participer, d'une façon ou d'une autre, le plus d'étudiants et de professeurs possible. On peut alors, grâce à cette démarche d'élaboration d'un programme d'évaluation de l'enseignement, sensibiliser les uns et les autres aux dimensions d'un bon enseignement, amorcer un dialogue: les étudiants comprendront peut-être mieux les difficultés que peuvent rencontrer les professeurs et ces derniers seront plus conscients des attentes des étudiants.

Que fera-t-on des résultats de l'évaluation?

C'est une autre décision à prendre très importante car on peut s'attendre à rencontrer beaucoup de difficultés si on utilise les résultats de l'évaluation à d'autres fins que celles auxquelles elles devaient servir. Il est nécessaire d'avoir une attitude ouverte en ce qui concerne l'utilisation des résultats, sinon ce sera l'objet de conflits.

Il s'agit, entre autres, de préciser ceci: qui aura accès aux résultats de l'évaluation, *qui évaluera*, c'est-à-dire qui portera le jugement après avoir examiné les données recueillies. Par exemple, les résultats peuvent être connus du professeur seulement ou du professeur et d'un conseiller pédagogique dans le cas d'une auto-évaluation, d'une évaluation qui vise avant tout à renseigner le professeur qui cherche à se perfectionner. S'il y a des incidences contractuelles (tout est fonction du but de l'évaluation), les résultats peuvent être transmis à l'administration et versés au dossier du professeur. Dans le cas d'une évaluation menée à des fins d'orientation des nouveaux étudiants, les résultats sont habituellement connus de tous les étudiants et professeurs.

Il est nécessaire que l'on sache précisément à *qui revient la prise de décision*, une fois que l'on a en main les résultats de l'évaluation. Il faut que l'utilisation qu'on veut en faire corresponde au but poursuivi par l'évaluation: dans le cas contraire (ce qui se produit malheureusement trop souvent), on suscite de nombreuses résistances dans le milieu. La réponse apportée à «que fera-t-on des résultats?» permet de réviser celle qu'on a donnée à «quel est le but de l'évaluation?», de façon à ce qu'il y ait cohérence entre ce qu'on veut faire et ce que l'on fera.

Il faut cependant rappeler que les *résultats doivent être interprétés en fonction des buts de l'évaluation, en tenant compte de la situation*, c'est-à-dire des conditions matérielles, du nombre d'étudiants et le reste. Nous reviendrons d'ailleurs sur ce point.

PROGRAMME D'ÉVALUATION DE L'ENSEIGNEMENT

Précisons les diverses *étapes qui composent un programme d'évaluation de l'enseignement*, c'est-à-dire étudions les questions qu'on doit se poser lors de l'élaboration d'un tel programme.

Quel est l'objet de l'évaluation?
Ce qu'on cherche à évaluer afin d'atteindre le but visé par l'évaluation est souvent confus ou imprécis, et pourtant des distinctions précises à ce sujet s'imposent. Par exemple, on peut distinguer les composantes du cours des caractéristiques du professeur.

Que cherche-t-on à évaluer au juste? La réponse à cette question détermine en grande partie l'utilité et la validité des résultats.

Les objets les plus fréquemment visés par une évaluation, parce qu'ils sont plus facilement mesurables peut-être, sont les suivants:

- La pertinence ou l'efficacité des méthodes pédagogiques utilisées pour animer un cours;
- la qualité ou l'utilité des méthodes utilisées pour évaluer les progrès de l'étudiant;
- les caractéristiques du professeur (volubilité, contact avec les étudiants, dynamisme, etc.);
- l'organisation du cours;
- le progrès de l'étudiant en ce qui a trait aux objectifs du cours à atteindre;
- les modifications d'attitudes chez les étudiants, à l'égard de la matière enseignée par exemple;
- la pertinence d'un cours et de son contenu par rapport à l'ensemble d'un programme de formation ou d'éducation;
- la pertinence d'un programme par rapport aux besoins de la clientèle étudiante, aux objectifs de ce programme, au marché du travail, et l'agencement de ses diverses composantes.

(Les deux derniers points font l'objet d'une évaluation que l'on qualifie habituellement «de programme»).

La spécification de ce qu'on veut mesurer préside au choix des instruments appropriés. Le choix des critères sera une explicitation de l'objet de l'évaluation.

Quels sont les objectifs de l'enseignement?
Les objectifs du cours ou du programme doivent être précisés lorsqu'on veut évaluer la pertinence de ces objectifs ou savoir s'ils ont été atteints.

De plus, le choix d'une méthode d'enseignement, le choix de modalités d'évaluation ne sont pas bons ou mauvais en soi. Leur pertinence dépend du contexte éducatif de même que des objectifs. L'efficacité d'un enseignement réside dans les relations entre ces divers éléments.

Quelles sont les sources d'information?
Dans plusieurs cas, il peut être difficile d'obtenir des mesures objectives pour évaluer les divers objets mentionnés au premier point. Il est assez fréquent que l'on demande alors aux étudiants, aux collègues ou au professeur lui-même de fournir ces mesures qui seront plus ou moins objectives selon qu'on se réfère à des faits ou à des opinions.

Les instruments pourront être différents selon qu'ils s'adressent au professeur lui-même, aux étudiants ou à ses collègues. Il faut évidemment associer les sources de données aux buts et aux objets de l'évaluation. Dans certains cas, on peut choisir plus d'une source d'information; ceci est préférable mais demande plus d'efforts.

Évaluation par les étudiants
On demande aux étudiants de fournir des données relatives à des critères reliés à l'enseignement et au professeur. Il est très fréquent que l'on s'adresse à eux: on considère que les étudiants sont les plus qualifiés pour évaluer l'enseignement (si des mesures plus objectives ne peuvent être disponibles) car ce sont eux qui entrent le plus souvent en contact avec le professeur. Il faut souligner que cette pratique facilite le dialogue professeur/étudiants. C'est aussi un signe tangible de l'importance accordée aux étudiants.

Mentionnons que l'on peut aussi demander aux anciens étudiants, aux diplômés, de donner leur avis. Cependant, les recherches de Centra, de Drucker et Remmers, nous incitent à croire que ceux-ci perçoivent l'enseignement reçu à peu près de la même façon que les étudiants actuels. C'est différent dans le cas de l'évaluation d'un programme: comme les diplômés connaissent mieux le milieu de travail, ils sont plus en mesure d'évaluer la pertinence du programme de formation.

Évaluation par les collègues
Les collègues peuvent peut-être plus difficilement dire ce qui se passe dans la classe d'un autre professeur, mais ils peuvent donner des indications quant à la participation à la vie du département pour autant que l'on considère cela comme faisant partie du rôle d'un professeur (le consensus n'est pas facile à faire sur ce point). Il faut aussi définir ce qu'on entend par collègues: s'agit-il des professeurs du même département, de ceux qui enseignent la même matière, de ceux qui travaillent dans une même équipe? Le chef de département est-il un collègue? La participation du chef de département est-elle prépondérante?

Les données recueillies auprès des collègues suscitent des problèmes surtout lorsque l'évaluation a des incidences contractuelles ou monétaires. Cela peut entraîner des jeux de rivalité, ou encore il n'est pas rare de voir un professeur refuser de juger un de ses pairs, de voir un syndiqué refuser de juger un autre syndiqué.

Même en ce qui a trait à une évaluation ayant pour unique but d'améliorer l'enseignement, l'évaluation par des collègues n'est pas sans difficultés. Par exemple, si on demande à un collègue de donner de l'information sur ce qui se passe en classe, une ou deux visites sont bien peu pour se former une opinion. Et il s'agit de l'avis d'une seule personne en plus. Dans le cas où on lui demande ce qu'il pense d'un plan de cours, cela peut sembler trop restreint comparativement à l'ensemble des activités pédagogiques. En dernier lieu, disons que s'il était possible de créer un climat sain dans lequel on puisse discuter des améliorations possibles à apporter à un enseignement, cela serait profitable autant pour les nouveaux professeurs que pour tout professeur qui révise son enseignement.

Auto-évaluation du professeur

Il y a diverses façons de pratiquer l'auto-évaluation. Le professeur peut avoir recours au même instrument que celui utilisé par ses collègues et ses étudiants et comparer ses données à celles recueillies auprès de ses collègues et de ses étudiants. On peut aussi enregistrer son enseignement sur ruban magnétoscopique[2] et il peut analyser, seul ou avec une personne-ressource, sa relation avec les étudiants en classe. Cette façon de procéder débouche plus facilement sur le perfectionnement car elle éveille moins la méfiance du professeur, il se crée moins de mécanismes de défense comme la négation des résultats. Evidemment, il s'agit d'un examen de conscience qui peut être bénéfique pour tous.

Pour ce qui est d'évaluer si un programme est adéquat, si ses composantes sont bien agencées, si les conditions matérielles sont appropriées, on peut demander l'avis des professeurs, des étudiants, des diplômés et des divers corps intermédiaires.

En terminant, il est à noter qu'aux informations fournies soit par les étudiants, les collègues, le professeur lui-même ou par une autre source doivent faire suite des décisions prises soit par le professeur, les étudiants ou les autorités, et ce, en fonction du but poursuivi par l'évaluation.

Quel type d'instrument choisir?

Après avoir franchi ces diverses étapes, il est possible de déterminer le type d'instrument qui serait le plus satisfaisant. Il existe plusieurs types d'instruments pour évaluer l'enseignement et il y en aurait sûrement beaucoup d'autres à explorer, à inventer. Nous n'exposerons ici que les types d'instruments les plus utilisés.

Un des types d'instrument les plus fréquemment employés est *le questionnaire*. Dans ce cas, il faut d'abord choisir entre un *instrument à questions fermées*, c'est-à-dire composé d'énoncés suivis d'un choix de réponses, et un *instrument à questions ouvertes* dans lequel aucune réponse n'est suggérée au répondant.

Un instrument à questions fermées est avantageux lorsqu'on désire que tous les étudiants se prononcent sur les mêmes points. On obtient de l'information sur des éléments précis. Lorsqu'on s'adresse à une population nombreuse, il est plus facile de procéder à la compilation et à l'analyse des réponses. Cependant, il ne faut pas croire que la compilation des réponses par l'ordinateur assure l'objectivité des répondants, surtout lorsqu'on leur demande de juger plutôt que de décrire.

Comme on désire obtenir de *l'information sur des points précis*, il faut que les énoncés soient clairs, simples, aussi brefs que possible, et ils doivent se référer à un seul trait ou comportement. Il faut bien faire attention à la rédaction des énoncés de telle sorte que l'information recueillie soit utilisable.

Voici quelques exemples d'énoncés d'un instrument à questions fermées:

- Les objectifs du cours sont clairs.
- Le professeur est disponible pour donner des consultations en dehors des heures de cours.
- La bibliographie fournie pour ce cours est adéquate.
- Et ainsi de suite...

On peut aussi trouver des questions plutôt que des énoncés dans un tel questionnaire.

Les instruments à questions ouvertes sont généralement composés de questions globales. Ils favorisent *la libre expression des répondants*: ceux-ci sont moins cantonnés dans l'évaluation de certains éléments de l'enseignement et des problèmes auxquels on n'aurait pas pensé peuvent ainsi être identifiés.

Par exemple, la technique des incidents critiques[3] utilise des questions ouvertes afin de recueillir des descriptions de situation ou d'événement précis indiquant des éléments à améliorer ou à conserver. Voici des exemples de questions ouvertes reliées à cette technique appliquée à l'évaluation d'un cours:

Essayez de vous rappeler un fait, une situation ou un événement spécifique que vous avez vu ou vécu au cours de la session et qui vous donne l'impression que certaines choses auraient besoin d'être améliorées dans ce cours. Ne parlez que d'un seul fait, situation ou événement sur cette feuille de réponse.

- Quand ce fait, cette situation ou cet événement s'est-il produit? --------
  ---

- A quel endroit ou dans quelles circonstances cela s'est-il produit? ------
  ---
- Qu'est-il exactement arrivé? ---

Cette technique permet d'identifier plusieurs points à améliorer ou à conserver, ceux-ci étant articulés sur des événements vus ou vécus.

Cependant, les réponses relatives à des instruments composés de questions ouvertes sont plus difficiles à analyser. Tous les étudiants ne s'étant pas prononcés sur les mêmes points, on ne peut en retour se prononcer sur l'importance des problèmes identifiés. Ces questionnaires peuvent être utilisés comme premier sondage, à la suite de quoi on pourra élaborer un second instrument, à questions fermées par exemple.

On essaie parfois de *combiner les deux types de questions ouvertes et fermées*. Par exemple, à la suite de questions fermées, on demande d'inscrire des suggestions, des commentaires, des points forts ou faibles du cours et le reste. Nous ne croyons cependant pas que cette façon de procéder permette de conserver tous les avantages des deux types de questions: les répondants sont orientés par les questions fermées.

Pour ce qui est d'un instrument à questions fermées, d'autres éléments sont aussi à considérer.

a) Il faut décider du *nombre d'alternatives* par énoncé. Ce nombre varie entre deux et dix et il est habituellement de quatre ou cinq.

b) Il faut habituellement décider du *genre d'échelle* à adopter, sauf lorsqu'on présente des alternatives n'ayant aucun lien entre elles. (Par exemple: «Quelle est la raison habituelle de vos absences au cours?» et les alternatives seraient les suivantes: 1. La maladie 2. Le manque d'intérêt du cours Etc.).

Nous exposons ici deux types d'échelle:

- Des chiffres sont associés aux diverses alternatives présentées pour chaque énoncé d'un questionnaire.

Par exemple:

1. totalement en désaccord
2. partiellement en désaccord
3. partiellement d'accord
4. totalement d'accord

Le répondant choisit alors, pour chaque énoncé, l'alternative qui représente sa position. Le poids «1» est alors donné à la première alternative et ainsi de suite.

- Un énoncé est suivi d'une ligne continue graduée qui indique des catégories en divers points du continuum.

Par exemple: 1 — 2 — 3 — 4 — 5

tout à fait en désaccord — tout à fait d'accord

A chaque énoncé, le répondant inscrit ou encercle le chiffre qui correspond à sa position.

(Notons que la première échelle n'a pas de point neutre tandis que la seconde en a un.)

c) Il faut décider du *mode d'inscription des réponses* et donner des directives claires quant à la façon de répondre. Ceci facilite, entre autres, la compilation des réponses.

Nous avons traité plus longuement de l'instrument composé de questions fermées. Il ne faudrait pas en déduire que cet instrument est plus avantageux. Il comporte tout simplement des étapes, des décisions supplémentaires.

Passons maintenant à des instruments autres que les questionnaires: *les entrevues individuelles ou de groupe* qui sont assez fréquemment utilisées. Les entrevues possèdent habituellement un avantage similaire à celui dont nous avons parlé dans la section sur les instruments à questions ouvertes: elles facilitent l'identification de problèmes insoupçonnés et permettent d'étayer ceux-ci en procédant au relevé de certains faits. Par une série de questions et de réponses, on peut cerner un problème et ses causes.

Les entrevues individuelles peuvent être utilisées, par exemple, pour émettre des hypothèses quant à la véritable raison d'un grand nombre d'abandons dans un cours que les étudiants réussissent par ailleurs assez bien, ou encore quant aux causes d'un climat de classe malsain. Grâce à cette façon de procéder, un répondant ne se laisse pas influencer par un autre. Cependant, c'est une opération coûteuse au point de vue temps; lorsqu'on a plusieurs répondants, on doit procéder par échantillonnage.

L'entrevue de groupe peut être utilisée pour évaluer un cours, un programme. Par exemple, si on a affaire à une centaine d'étudiants-répondants, on peut les diviser en sous-groupes d'une douzaine d'individus assistés d'un animateur. Les données recueillies dans ces sous-groupes peuvent ensuite faire l'objet d'une plénière. On peut ainsi vérifier s'il y a consensus: la majorité des étudiants sont d'accord ou non avec l'identification de tel problème. Les entrevues de groupe facilitent le dialogue entre les divers intéressés et, lors de la plénière, on peut obtenir immédiatement du feedback. Par exemple, il est possible pour le(s) professeur(s) ou le directeur de réagir aux données recueillies, de discuter tout de suite des modifications à apporter au cours ou au programme permettant de résoudre certains des problèmes identifiés.

Selon les cas, un animateur qui n'est pas du «milieu» peut mener l'entrevue, et ce, à l'avantage des participants. Il y a alors moins de possibilités de donner des réponses biaisées. Cependant, afin de susciter les échanges entre les diverses parties concernées, on peut les faire toutes participer aux entrevues de groupe. Par exemple, le professeur peut discuter avec ses étudiants de son enseignement; les professeurs, les autorités et les étudiants peuvent participer à des discussions portant sur l'évaluation du programme.

En plus de choisir un animateur et les participants, on doit aussi préciser le degré de structuration des entrevues. On peut demander quelque chose d'aussi global que d'identifier les problèmes, et éventuellement leurs causes, ou encore préparer des séries de questions inhérentes aux diverses facettes d'un cours ou d'un programme: par exemple, l'évaluation de l'apprentissage, les méthodes d'enseignement et le reste. Le choix dépend de ce qui est visé par le programme d'évaluation de l'enseignement. Ainsi, une entrevue a avantage à être moins structurée lorsqu'on ne peut émettre d'hypothèses sur un problème et ses causes. Une entrevue struturée permet par ailleurs de vérifier divers points précis, de soumettre à la critique certaines hypothèses. Une entrevue structurée facilite l'analyse des données, de même que la tenue des plénières, s'il y a lieu, car on peut plus aisément avoir une idée des consensus, de l'ordre d'importance des problèmes.

Les entrevues individuelles ou de groupe sont aussi utilisées avant ou après un autre instrument d'évaluation. Par exemple, on peut discuter avec les répondants à partir de l'analyse de leurs réponses à un questionnaire, ou immédiatement après qu'ils ont rempli individuellement un questionnaire. Les entrevues peuvent aussi constituer un premier pas dans un processus d'évaluation de l'enseignement en donnant un aperçu de la situation, un instrument plus structuré pouvant y faire suite.

Notons que lorsque le nombre de répondants est peu élevé, il peut être plus avantageux de discuter avec ceux-ci: leurs réponses à un questionnaire, même sophistiqué, n'étant pas fiables.

Parmi les autres instruments qu'il est possible d'utiliser, nous avons déjà mentionné *l'analyse interactionnelle*[4] lorsqu'on désire analyser les interactions professeur/étudiants en classe. Par exemple, on peut évaluer le temps de parole accordé aux étudiants en classe, comment le professeur pose des questions et ainsi de suite. Un tel instrument ne s'applique donc pas à toutes les stratégies d'enseignement: par exemple, il ne s'applique pas à l'enseignement individualisé.

Quels sont les critères?

Cette question ne se pose pas en ce qui concerne tous les instruments. Par exemple, elle ne se pose pas lorsqu'on utilise la technique des incidents critiques, mais on ne peut la négliger dans le cas d'un instrument structuré (exemple: un instrument composé de questions fermées). Les critères d'évaluation sont l'explicitation de l'objet à évaluer et serviront à formuler des questions, des énoncés.

Les critères d'évaluation varient selon la réponse que l'on donne aux questions précédentes, c'est-à-dire que le choix des critères peut dépendre du but de l'évaluation, de l'objet mesuré, de la source d'information et ainsi de suite.

Le cas le plus courant est celui où *l'étudiant* sert de source d'information pour évaluer l'enseignement d'un professeur. Il est habituellement admis que les critères doivent autant que possible porter sur des segments définis de comportements observables.

L'analyse factorielle permet de découvrir les facteurs (ou composantes de base) sous-jacents au choix des professeurs considérés comme bons ou mauvais. Ces facteurs varient quelque peu selon les instruments analysés. Une autre façon d'établir une liste des diverses dimensions d'un bon enseignement consiste à demander aux professeurs et aux étudiants d'en dresser une liste et, s'il y a lieu, d'indiquer celles qui leur semblent les plus importantes. Il y a évidemment plusieurs critères possibles. En voici quelques exemples:

Atmosphère de la classe:
- respect des étudiants,
- climat de motivation,
- confiance dans la réussite,
- ouverture aux diverses idées,
- capacité d'adaptation.

Professeur:
- disponible,
- dynamique,
- tolérant,
- communicatif,
- compétent,
- intéressé à sa matière,
- bien préparé,
- honnête intellectuellement.

Cours :

- bien organisé et structuré,
- objectifs clairement définis,
- objectifs réalistes,
- utilisation de méthodes adéquates,
- évaluation juste,
- applications concrètes,
- travaux à faire: quantité, pertinence, etc.
- participation étudiante,
- remise des travaux à bref délai,
- feedback adéquat.

Il faut cependant que ce soient des critères que les étudiants puissent juger à partir de leur capacité d'observation. Par exemple, il est peut-être difficile, pour un très jeune étudiant, de juger si un professeur l'aide à mieux structurer sa pensée ou s'il a contribué à augmenter sa capacité de porter un jugement. De plus, il faut se demander si on a choisi des critères pertinents et utiles: par exemple, «avoir le sens de l'humour» n'est peut-être pas essentiel pour enseigner d'une façon efficace. Il est à noter que les recherches de Constin, de Ranan, de Gustad et de Kent n'indiquent que des relations peu convaincantes entre la personnalité d'un professeur et son efficacité. Evidemment, si un professeur désire savoir ce que ses étudiants pensent de son «sens de l'humour», il peut leur poser la question ou faire l'analyse d'un ruban magnétoscopique enregistrée en classe[5].

Lorsque les *collègues* alimentent la source de données, les critères peuvent être les mêmes que ceux qui sont utilisés par les étudiants lorsque ces professeurs vont en classe à titre d'observateurs par exemple. Le collègue peut aussi être amené à donner son avis après avoir examiné un plan de cours, des notes de cours, des examens et le reste.

On peut lui demander de juger la participation du professeur à la vie du département en fonction de critères comme l'esprit de collaboration, la participation à des comités, la participation à des discussions entre collègues sur l'enseignement et ainsi de suite. Parfois, on lui demande si le professeur est au courant des développements dans son domaine (ou encore, dans celui de ses collègues), s'il s'informe et s'intéresse aux recherches de ses collègues, si sa recherche est originale et le reste.

On peut, dans certains cas, recourir à des mesures moins subjectives comme *l'évaluation par objectifs*. Il s'agit alors de vérifier si les étudiants ont atteint les objectifs relatifs à chaque cours. Ceci permet d'évaluer l'efficacité de l'enseignement, de préciser si les objectifs sont valables, s'ils doivent être modifiés ou rejetés. Habituellement, cette façon de faire a pour but d'améliorer l'enseignement et non de juger le professeur. On cherche les causes des écarts entre les attentes et le résultat obtenu. Une telle évaluation demande un certain effort car elle suppose qu'on précise davantage les objectifs et qu'on mette au point des moyens d'évaluation qui visent à mesurer l'atteinte de ces objectifs.

L'évaluation des changements d'attitudes chez les étudiants peut être des plus importantes dans l'évaluation de l'enseignement. Ainsi, l'acquisition des connaissances au détriment d'une certaine motivation n'est pas souhaitable. C'est d'autant plus important lorsqu'un changement d'attitudes constitue un des objectifs du cours (parfois même c'est l'objectif prioritaire).

Il existe des mesures moins subjectives que celle qui consiste à demander à l'étudiant s'il est plus motivé (ou non) à la fin de son cours qu'au début. Par exemple, on peut obtenir certaines informations en employant des mesures indirectes: demander à l'étudiant s'il a suivi ce cours seulement parce qu'il était obligatoire, s'il a suivi un cours non obligatoire qui y fait suite et le reste. On peut aussi considérer comme critères le nombre de lectures et d'activités non obligatoires auxquelles s'adonnent les étudiants. Lorsqu'on veut modifier les attitudes des étudiants par rapport à un objet précis (la prévention, par exemple), on précise d'abord celui-ci, puis on peut construire des instruments permettant de mesurer les attitudes par rapport à des énoncés relatifs à cet objet, énoncés sur lesquels les étudiants se prononcent.

Nous n'avons donné ici que quelques exemples de critères. Or, une situation donnée peut suggérer des critères qui n'ont pas été mentionnés. Notons cependant qu'il est avantageux de consulter les répondants sur les critères qu'ils jugent pertinents. Ceci permet de mieux déceler les problèmes en identifiant les critères appropriés.

Il n'est pas toujours facile d'obtenir un consensus sur le choix des critères, surtout lorsque les professeurs utilisent des stratégies d'enseignement fort variées. C'est pourquoi il peut être avantageux *d'évaluer chacun des enseignements sur mesure* en adaptant le choix des critères aux diverses composantes de l'enseignement. Il est alors possible de fournir une information répondant aux besoins de chacun des professeurs. Il faut cependant construire, dans ce cas, plusieurs instruments. Ceux-ci peuvent avoir une partie commune et une partie correspondant à chacune des stratégies d'enseignement. S'il est plus difficile alors de comparer les résultats des divers professeurs, à des fins de promotion ou de réengagement (évaluation sommative), on évite par contre, grâce à cette façon de procéder, les ennuis d'un questionnaire unique et global, ne pouvant guère renseigner, ou risquant d'enfermer les professeurs dans un modèle d'enseignement.

A noter: lorsque l'évaluation vise la promotion ou le réengagement des professeurs, il est fréquent que le nombre de critères soit plus restreint que lorsqu'il s'agit d'une évaluation formative de l'enseignement. Surtout si le professeur a modifié sa stratégie, il a besoin de nombreuses informations pour vérifier la qualité de son enseignement.

Construction ou choix d'instrument(s)

S'il existe déjà un ou des instruments répondant aux conditions requises, le temps de construction peut être épargné. Néanmoins, plusieurs préfèrent tout de même élaborer leur(s) instrument(s) afin de susciter la participation du milieu, de faire croître la motivation, de sensibiliser la communauté à la pédagogie et d'obtenir des instruments mieux adaptés aux besoins locaux.

Comme on l'a mentionné précédemment, il peut être avantageux d'avoir *plusieurs instruments pour évaluer l'enseignement, dépendant de la stratégie d'enseignement utilisée*. Il peut même avoir autant d'instruments que de professeurs, que de stratégies d'enseignement différentes, et, pourquoi pas, un instrument composé d'un tronc commun et d'une partie adaptable. Par exemple, cela va de soi, dans un instrument à questions fermées, un énoncé portant sur la clarté des exposés ne peut être pertinent pour un professeur qui n'utilise pas l'exposé. S'il expérimente une nouvelle méthodologie, il est possible qu'il soit, par ailleurs, intéressé à des questions très précises qui ne peuvent être utilisées par aucun autre professeur de son département: on aboutit alors au questionnaire sur mesure.

Il est important de souligner *qu'on ne s'improvise pas constructeur d'instrument d'évaluation, rédacteur de questions*. La consultation d'un spécialiste permet d'éviter de nombreuses erreurs. Par exemple, des énoncés imprécis comme ceux-ci ne sauraient figurer dans un instrument à questions fermées: «Votre professeur est compétent.», «Votre professeur est un bon pédagogue.», «Votre professeur est disponible.» et ainsi de suite. Dans le dernier énoncé, entre autres, on n'a pas précisé s'il s'agissait de disponibilité relative à des consultations qui se donnent en dehors des périodes de classe.

Il faudra *pré-expérimenter les instruments*. En ce qui a trait à un questionnaire ou à une entrevue structurée par exemple, on vérifie si les répondants comprennent bien les questions. On peut aussi vérifier les qualités métrologiques des instruments. S'il s'agit d'un instrument composé de questions, on peut aussi voir s'il y a trop de questions: il faut éviter les questionnaires trop longs, qui découragent les répondants. Un questionnaire succinct est plus séduisant; la présentation visuelle est importante. Rappelons tout de même que lorsqu'il s'agit d'évaluation formative, il est normal de trouver un peu plus de questions: on désire souvent une information plus détaillée pour améliorer l'enseignement. L'expérimentation des instruments permet donc d'estimer le temps d'administration, de vérifier si les directives et la présentation de l'instrument sont adéquates.

Il faut se rappeler que les divers instruments ne s'excluent pas mutuellement. On peut en utiliser plusieurs dépendant des sources d'information (étudiants, collègues, etc.) auxquelles on s'adresse. Et on peut en utiliser de divers types, et ce pour une même source d'information (par exemple, un questionnaire suivi d'une entrevue de groupe). Evidemment, lorsqu'on utilise plusieurs instruments et qu'on procède à une évaluation sommative, il est possible que l'on doive préciser s'ils ont tous un même poids, sinon on doit déterminer le poids de chacun.

Quelles seront les modalités?

On doit aussi préciser les modalités d'administration de l'instrument d'évaluation. Voici certaines questions qu'on peut se poser à ce sujet:

- Quelle doit être la fréquence de cette évaluation?
  (Il faut penser que les répondants - étudiants, collègues ou autres - peuvent se désintéresser de l'opération s'il y a réévaluation, de session en session, sans qu'il y ait vraiment de changements dans le cours, dans le programme, et conséquemment de modifications aux instruments d'évaluation.)
- A quel moment l'évaluation doit-elle avoir lieu pour être efficace?
  (Il faut fixer une période durant la session ou l'année pour administrer l'évaluation: dans certains cas, on opte pour la mi-session afin d'améliorer l'enseignement, d'engager le dialogue; dans d'autres, on préfère la fin de la session s'il ne peut y avoir de modifications substantielles en cours de route ou lorsque les étudiants/répondants veulent se protéger contre toutes représailles éventuelles.)
- Dans quelles conditions administrera-t-on les instruments?
  (Il faut estimer le temps d'administration et réserver une période suffisante pour permettre aux répondants d'évaluer en prenant bien leur temps. Il faut ainsi éviter de demander à des étudiants de répondre à un questionnaire pendant leur heure de lunch, après un cours et le reste: il faut leur accorder une période particulière.)
- L'évaluation doit-elle être obligatoire ou les professeurs pourront-ils choisir ou non de profiter des instruments mis à leur disposition?
  (On comprendra que la réponse peut varier, tout est fonction du but de l'évaluation.)
- Est-ce que tous les professeurs et/ou tous les cours seront évalués?
  (Même lorsque l'évaluation est obligatoire, elle ne l'est parfois que pour les professeurs en probation.)
- Est-ce que l'anonymat des répondants sera assuré?
  (Lorsque les répondants sont des étudiants, la tendance est de leur assurer cet anonymat; ce qui n'est pas toujours le cas lorsque les répondants sont des collègues.)
- Est-ce que d'autres variables (âge, sexe, orientation de l'étudiant, années d'expérience et titre du professeur, etc.) seront considérées?
  (Ces variables sont importantes si on veut faire des recherches sur les liens entre ces variables et les réponses données à l'instrument, et ce, afin de pouvoir donner une meilleure interprétation des résultats.)
- S'il s'agit d'un instrument administré aux étudiants, par qui doit-il être administré?
  (Il vaut parfois mieux que ce soit un tiers, le professeur étant absent de la classe à ce moment-là; les recherches ont montré que, sans cela, les réponses sont biaisées positivement.)
- S'il s'agit d'un questionnaire, est-il préférable d'envoyer celui-ci par la poste aux gens concernés?
  (S'il peut être avantageux, du point de vue de la motivation, de n'exercer aucune pression pour que le questionnaire soit rempli à tel moment plutôt que tel autre, il faut aussi tenir compte du taux de répondants qui risque alors d'être plus faible.)

- Est-ce qu'on a pris tous les moyens possibles pour déceler les problèmes? (Par exemple, lorsqu'on administre en fin de session un instrument aux étudiants, ceux qui ont abandonné le cours n'ont pu s'exprimer et, dans certains cas, il pourrait être d'une grande importance de savoir pourquoi ceux-ci ont abandonné et ce qu'ils pensent du cours.)

Comment présenter les résultats?

En choisissant ou en construisant l'instrument, il faut penser au type de résultats à présenter et à la façon de les présenter. Les résultats que l'on présente dépendent en majeure partie du type d'instrument d'évaluation choisi. Par exemple, si on a utilisé l'entrevue de groupe, on peut trouver dans un rapport une description de la population, la retranscription des réponses de chacun des groupes, les consensus relevés à la plénière, puis l'analyse de ces réponses. Pour ce qui est d'un questionnaire fermé, outre la description de la population, on peut y trouver la moyenne, la distribution des réponses, l'écart type, ou autre statistique calculée en fonction de chacune des questions. On doit faire le choix des statistiques à présenter: le but de l'évaluation est le facteur à considérer dans ce choix. Par exemple, lorsqu'on établit des comparaisons entre professeurs (évaluation normative), on doit choisir si on compare le professeur à un groupe de professeurs qui enseigne la même matière que lui, aux professeurs de son département ou encore à l'ensemble des professeurs. La réponse variera selon ce qu'on entend faire des résultats. Il faut aussi choisir si on présentera des percentiles, des stanines ou autres statistiques se référant au groupe. Notons cependant qu'il faut éviter les cotes uniques représentant le résultat d'une personne concernant les multiples dimensions d'un bon enseignement.

On doit aussi penser, à ce moment-ci ou après l'analyse des résultats, à la présentation visuelle des résultats. Par exemple, utilisera-t-on des graphiques pour illustrer la distribution des réponses à une question? Il est important, pour ce qui est de l'impact d'une telle opération, de présenter les résultats de façon précise et concise.

Quels sont les coûts?

Il faut tenir compte du temps et des sommes nécessaires à la construction et à l'administration des instruments d'évaluation de l'enseignement: le matériel requis, le temps de l'ordinateur, les salaires d'une ou de plusieurs personnes engagées dans ce programme et le reste.

EXPÉRIMENTATION ET SUITES

Après avoir élaboré une politique et ensuite un programme d'évaluation de l'enseignement, nous arrivons à *la phase de l'expérimentation et de ses suites*.

L'expérimentation

Il s'agit ici de procéder à l'évaluation de l'enseignement au moyen des instruments élaborés à cette fin, en respectant les modalités d'administration. Si on s'est bien assuré de l'intérêt du milieu pour ce programme d'évaluation, on ne devrait pas rencontrer de problèmes majeurs au cours de cette phase, surtout si on peut garantir de bonnes modalités d'administration. Il est important de rappeler que des modalités d'administration inadéquates peuvent affaiblir un programme d'évaluation.

L'expérimentation permet d'accumuler des données pour évaluer l'enseignement et évaluer le processus d'évaluation de l'enseignement.

L'analyse des résultats

On compile et on analyse les données recueillies lors de l'application du programme d'évaluation de l'enseignement en fonction des types de résultats que l'on a choisi de présenter.

Il faut *être prudent dans l'interprétation des résultats*. Rappelons d'abord qu'il faut les interpréter en fonction des buts de l'évaluation et qu'il faut aussi tenir compte du contexte, c'est-à-dire des conditions matérielles, du nombre d'étudiants et le reste. Par exemple, le fait qu'il s'agisse d'un cours optionnel ou obligatoire semble un facteur à considérer lorsqu'on évalue l'enseignement: en effet, les études de Gage, de Lovell et Haner indiquent que les cours optionnels auraient tendance à recevoir des cotes plus élevées.

Même si on utilise des instruments très structurés, techniquement bien construits, il faut rappeler ceci: lorsque le nombre de répondants est trop faible, l'analyse des résultats peut être compromise. Ainsi, en ce qui a trait à un questionnaire fermé adressé à des étudiants, les études de Remmers et Guthrie indiquent que ce type d'évaluation est fidèle lorsqu'on a vingt-cinq répondants et plus. Par exemple, lorsque seulement une douzaine d'étudiants répondent à un tel questionnaire, leurs réponses pourraient être très différentes de celles récoltées auprès de douze autres étudiants évaluant le même objet. On comprend donc à quel point il faut être prudent dans l'analyse des résultats; il faut éviter de tirer des conclusions à partir de données qui ne sont pas fiables. On peut être encore plus exigeant lorsque les données sont utilisées à des fins sommatives puisque les retombées de l'évaluation ont un caractère plus définitif.

L'évaluation de l'enseignement et ses conséquences

L'analyse des résultats permet d'évaluer l'enseignement, de poser un jugement. Dans le cas d'une évaluation menée à des fins formatives, un professeur peut porter un jugement sur son enseignement à partir de l'analyse de ses résultats; il peut ensuite entreprendre, s'il y a lieu, de modifier ses modalités d'enseignement et d'évaluation, sa façon de communiquer avec les étudiants et le reste. Dans le cas d'une évaluation menée cette fois à des fins sommatives, les autorités décident, par exemple, d'accorder ou pas une promotion à un professeur, de resubventionner ou pas un programme (dans le cas d'une évaluation de programme).

Cependant, *une première évaluation de l'enseignement n'est souvent qu'indicative:* Il est nécessaire de procéder à d'autres évaluations de l'enseignement pour s'assurer des résultats, pour mieux diagnostiquer les problèmes. Par exemple, on accumulera les données échelonnées sur quelques sessions avant d'accorder ou pas une promotion.

L'évaluation du processus d'évaluation de l'enseignement

L'évaluation de l'enseignement est un processus dynamique. Comme on est souvent intéressé à réévaluer l'enseignement, on doit vérifier si le processus d'évaluation est adéquat, si on doit le corriger, le modifier.

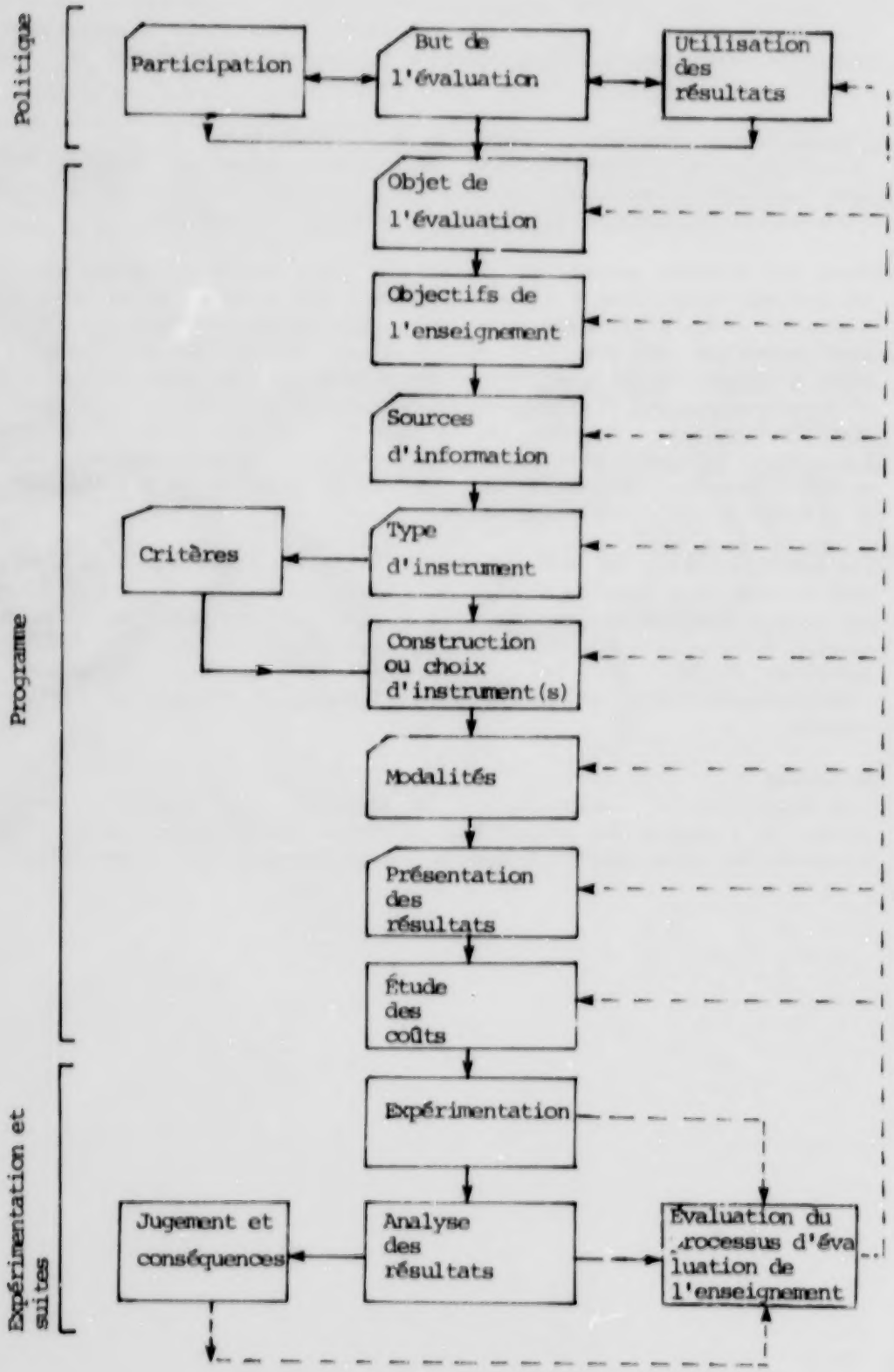
Parmi les diverses données qui peuvent servir à évaluer le processus, il y a les données recueillies à l'expérimentation: par exemple, si le temps d'administration prévu s'est avéré trop court, on fera les modifications qui s'imposent. L'analyse des résultats peut indiquer des pistes: par exemple, pour mieux diagnostiquer un problème, on devra construire un instrument plus précis en tenant compte des résultats de cette première évaluation. Il se peut aussi que l'on veuille vérifier (en analysant les résultats d'une première évaluation) l'impact des modifications apportées à un programme ou à un cours; on devra probablement apporter des changements au processus d'évaluation avant de procéder à une seconde évaluation.

Les modifications qu'on peut apporter au processus d'évaluation de l'enseignement peuvent être majeures ou mineures. Elles peuvent toucher une ou plusieurs des étapes du processus: par exemple, on peut modifier le but de l'évaluation, changer d'instrument d'évaluation ou y apporter des corrections, modifier les modalités et le reste. On comprend donc qu'un processus d'évaluation de l'enseignement n'est pas inaltérable, il faut le modifier en fonction des besoins.

En résumé

Le diagramme qui suit résume la démarche pour évaluer l'enseignement. On y trouve les étapes pour élaborer une politique et un programme d'évaluation, ainsi que les étapes de l'expérimentation et de ses suites.

DÉMARCHE POUR ÉVALUER L'ENSEIGNEMENT



Conclusion

Ce texte ne prétend pas résoudre les problèmes qui surviennent lors d'une évaluation de l'enseignement. Il vise tout au plus à exposer une démarche possible pour évaluer l'enseignement, à prévenir des difficultés inhérentes à certaines étapes, à présenter quelques choix, quelques réponses possibles à des questions que pose l'évaluation de l'enseignement. Avant tout, il s'agissait d'exposer un processus général... indépendamment du but visé par l'évaluation, du type d'instrument choisi et ainsi de suite.

Notes et références bibliographiques

1.Les termes relatifs au type d'évaluation sont définis dans le texte de Fontaine, France, «Évaluer? Mesurer?» dans *Dossier sur l'évaluation*, pp. 5-9.

2.Dussault, Gilles et al., *L'analyse de l'enseignement*. Montréal, Les Presses de l'Université du Québec, 1973, 311 p..

3.Cyr, Jean-Marc, *La technique des incidents critiques*. Montréal, Service pédagogique, Université de Montréal, 1978, 4 p..

4.Dussault, Gilles et al., *L'analyse de l'enseignement*. Montréal, Les Presses de l'Université du Québec, 1973, 311 p..

5.Id., ibid.

Section 3

Des pratiques d'évaluation de l'apprentissage

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Les tests sur les prérequis et les pré-tests

France Fontaine

Le présent texte veut expliciter ce que sont les *tests sur les prérequis et les pré-tests*. Nous préciserons aussi quelles sont leur utilisation et leurs diverses modalités d'application.

Les tests sur les prérequis

Si tous les étudiants d'un cours donné ont préalablement suivi un cours exigé par le programme d'études comme prérequis, vous savez fort bien qu'ils n'en ont pas tous bénéficié également. Même s'ils avaient tous obtenus la moyenne de 70% pour ce cours prérequis, cette moyenne serait le fruit de savoirs différents. Les disparités peuvent être plus marquées si les étudiants proviennent d'institutions différentes.

Ce fait incite les professeurs à faire des révisions de matière au début de leur cours. Ils veulent ainsi rafraîchir la mémoire des étudiants afin de leur faciliter l'apprentissage de la matière nouvelle en comptant sur les acquis des années précédentes. Or, que se produit-il? Que la révision soit courte ou longue, elle ennuie certains tandis que pour d'autres, elle est insuffisante.

Un test sur les prérequis d'un cours est un test de dépistage des difficultés que peuvent éprouver les étudiants à l'égard des matières requises pour suivre ce cours.

Le but habituellement visé par ces tests est de donner du feedback à chaque étudiant pour qu'il puisse lui-même apporter des correctifs, s'il y a lieu. L'étudiant est responsable de sa démarche et sa note n'est en rien affectée par le résultat. Le professeur peut aussi s'en servir pour apporter une aide individualisée à ses étudiants ou axer sa révision sur les faiblesses les plus fréquemment observées sur l'ensemble des étudiants. Cette évaluation qui favorise surtout un meilleur apprentissage est dite formative. L'évaluation est

dite sommative si le test sur les prérequis vise à identifier les étudiants auxquels on imposera un cours d'appoint, par exemple. Le test sert alors au classement des étudiants.

Revenons à l'évaluation formative des prérequis. Si les objectifs du cours sont rédigés, il est plus facile de préciser les prérequis en termes de notions ou même d'objectifs préalables en vue d'atteindre les objectifs du cours. Une fois cette étape franchie, il faut produire un test portant sur ces notions ou objectifs prérequis.

On doit aussi concevoir et produire, s'il y a lieu, les moyens d'apprentissage pour chacun de ces objectifs prérequis. Ces moyens peuvent être des lectures, des exercices, des exposés enregistrés sur ruban magnétoscopique et le reste. L'étudiant pourra ainsi combler ses lacunes sans perdre de temps à réviser ce qu'il sait déjà.

Il est important d'indiquer à l'étudiant que telle notion (ou objectif) est préalable à l'étude de telle partie (ou objectifs) du cours et d'y faire référence au moment opportun durant la session. On peut inciter l'étudiant à faire toute sa révision au début du cours. On peut aussi trouver préférable qu'il la fasse au fur et à mesure des besoins constatés durant le cours. Il peut aussi être avantageux de préparer un second test par lequel l'étudiant vérifiera son travail par rapport à ses lacunes personnelles décelées au premier test. Il ne répond alors qu'aux questions portant sur les notions ou objectifs pour lesquels il a éprouvé des difficultés au premier test.

Nous donnerons ici des exemples d'analyse de prérequis pour deux situations différentes d'enseignement.

a) Considérons d'abord un stage en obstétrique-gynécologie au niveau de l'externat[1]. Nous avons choisi trois objectifs du stage pour lesquels nous indiquons les objectifs prérequis. Ces objectifs prérequis sont des objectifs de cours théoriques préalables au stage.

Objectif du stage	Objectifs prérequis
L'externe sera capable de faire le diagnostic de grossesse face à une patiente qui n'a pas été menstruée depuis deux mois.	L'étudiant sera capable: • d'énoncer les symptômes de grossesse; • de décrire les signes pertinents de grossesse.
L'externe sera capable de reconnaître, en présence d'une patiente, tous les facteurs permettant l'identification des grossesses et des fœtus à risque élevé.	L'étudiant sera capable: • d'énumérer les facteurs de risques identifiables à l'histoire de cas, à l'examen physique, par les examens de laboratoire pour le diagnostic de la maladie fœtale, par les examens de laboratoire pour le diagnostic de la maladie maternelle; • de donner l'explication de chacun des facteurs à risque.
L'externe sera capable, en présence d'un cas réel d'un abcès de la glande de Bartholin, de faire le diagnostic de la lésion.	L'étudiant sera capable de: • décrire les signes et les symptômes de l'abcès de la glande de Bartholin; • décrire les agents en cause; • décrire l'évolution de la maladie.

On imagine facilement la partie du test sur les prérequis qui se rapporterait à ces objectifs prérequis.

b) L'exemple suivant est tiré d'un deuxième cours de calcul différentiel et intégral de niveau collégial (Math 203). Au lieu d'indiquer, comme dans l'exemple précédent, les objectifs du cours et les objectifs prérequis, nous donnons en exemple une partie d'un auto-test sur les prérequis de ce cours.

Directives

Le but de ce texte est de vous aider à vous rappeler les notions de fonctions, limites, continuité et dérivée qui vous seront utiles dans ce cours de M-203. Il n'en tient qu'à vous de vous en servir ou pas.

Vous trouverez pour chacune de ces notions: le numéro du texte correspondant dans le volume de référence si vous avez besoin de vous rappeler la théorie en question, les notions et chapitres du cours Math 203 où ces notions vous seront utiles et des problèmes correspondants à chacune de ces quatre notions.

Un solutionnaire vous sera distribué pour que vous puissiez vérifier vous-même vos connaissances en ces domaines.

Vous pouvez faire votre révision avant d'aborder la matière de Math 203 ou encore ne la faire qu'au moment où vous abordez les notions où elles sont requises dans le cours, c'est-à-dire que vous pouvez la faire globalement ou par parties en cours de session.

Fonctions (texte 0-3)

- La question 1 sert dès le début (chap. 1) et pour les techniques d'intégration (chap. 11).
- La question 2 sert pour :
  les hypothèses des théorèmes d'analyse (chap. 3-4),
  le calcul de surface (chap. 10),
  les applications de l'intégrale (chap. 12).
- La question 3 sert pour les intégrales de fonctions composées (chap. 11).
- La question 4 sert pourles intégrales de fonctions composées (chap. 11).

1-Donnez un exemple de

a) fonction affine

b) fonction constante

c) fonction identité

d) fonction polynomiale

e) fonction quadratique

f) fonction rationnelle

g) fonction transcendante

2-Tracez le graphe de

a) $f : x \longrightarrow |x|$ (valeur absolue)

b) $f : x \longrightarrow [x]$ (valeur entière)

c) $f : x \longrightarrow \sin x$

f) $f : x \longrightarrow e^x$

g) $f : x \longrightarrow \ln^x$

h) $f : x \longrightarrow |x^2 - 1|$

d) $f : x \longrightarrow \cos x$

e) $f : x \longrightarrow \text{tg } x$

c') $f : x \longrightarrow x^2 + 1$

3- Trouvez $f \circ g$ et $D_{f \circ g}$ (domaine) où $f \circ g$ est la composition des fonctions f et g.

a) $f : x \longrightarrow 2x + 3$

$g : x \longrightarrow x^2 - 1$

b) $f : x \longrightarrow \sqrt{x - 2}$

$g : x \longrightarrow x^2 + 1$

4- Décomposez les fonctions suivantes en deux fonctions g et h.

a) $f : x \longrightarrow \sqrt[3]{3x^2 - 7}$

$f : x \longrightarrow (7x + 12)^{11}$

Limites (texte 0-5 et 0-7)

- Les limites de tout genre peuvent être utiles pour la règle de L'Hospital (chap. 4).
- Les limites à l'infini sont nécessaires pour la convergence des suites (chap. 5) et pour les séries dans le cas du test du rapport (chap. 6) et pour les intégrales impropres (chap. 11).

1. $\lim\limits_{x \to 0} \dfrac{4x^2 + 3x + 7}{x^3 + 2x - 6}$

2. $\lim\limits_{x \to 2} (3x - 5)^{14}$

Et ainsi de suite.

Les pré-tests

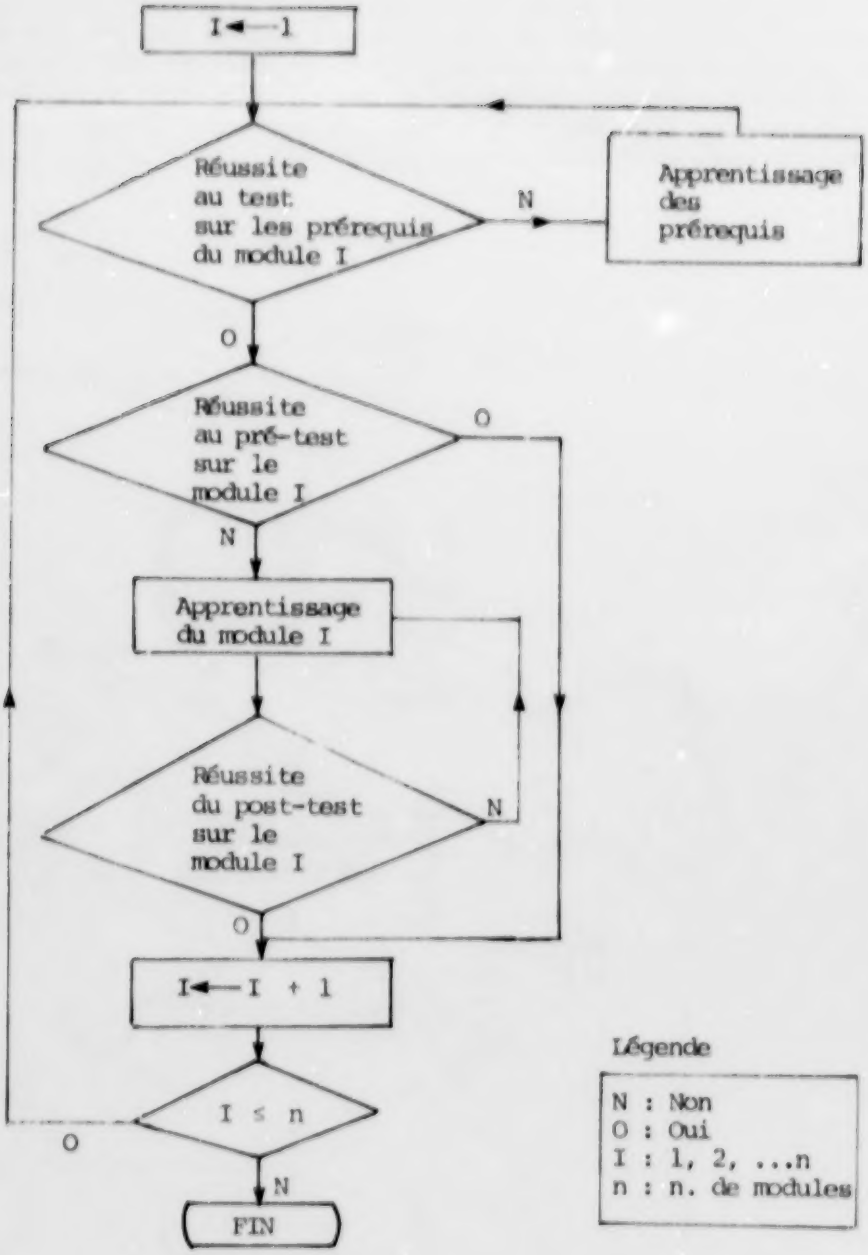
Les pré-tests sont utilisés dans un cours lorsque des étudiants, soit par des cours antérieurs, soit par culture personnelle, peuvent atteindre certains objectifs ou regroupement d'objectifs du cours sans formation supplémentaire.

Les pré-tests se sont développés dans le contexte de l'enseignement individualisé, qu'il s'agisse d'enseignement programmé ou modulaire. Ainsi, si un étudiant démontre, avant de recevoir la formation prévue dans le cours, qu'il atteint les objectifs de telle partie de cours, il passe à l'étape suivante. Sinon, il procède à l'apprentissage prévu, qui est ensuite évalué par un post-test. Le post-test est équivalent au pré-test mais il est appelé ainsi car il a lieu après certaines activités d'apprentissage.

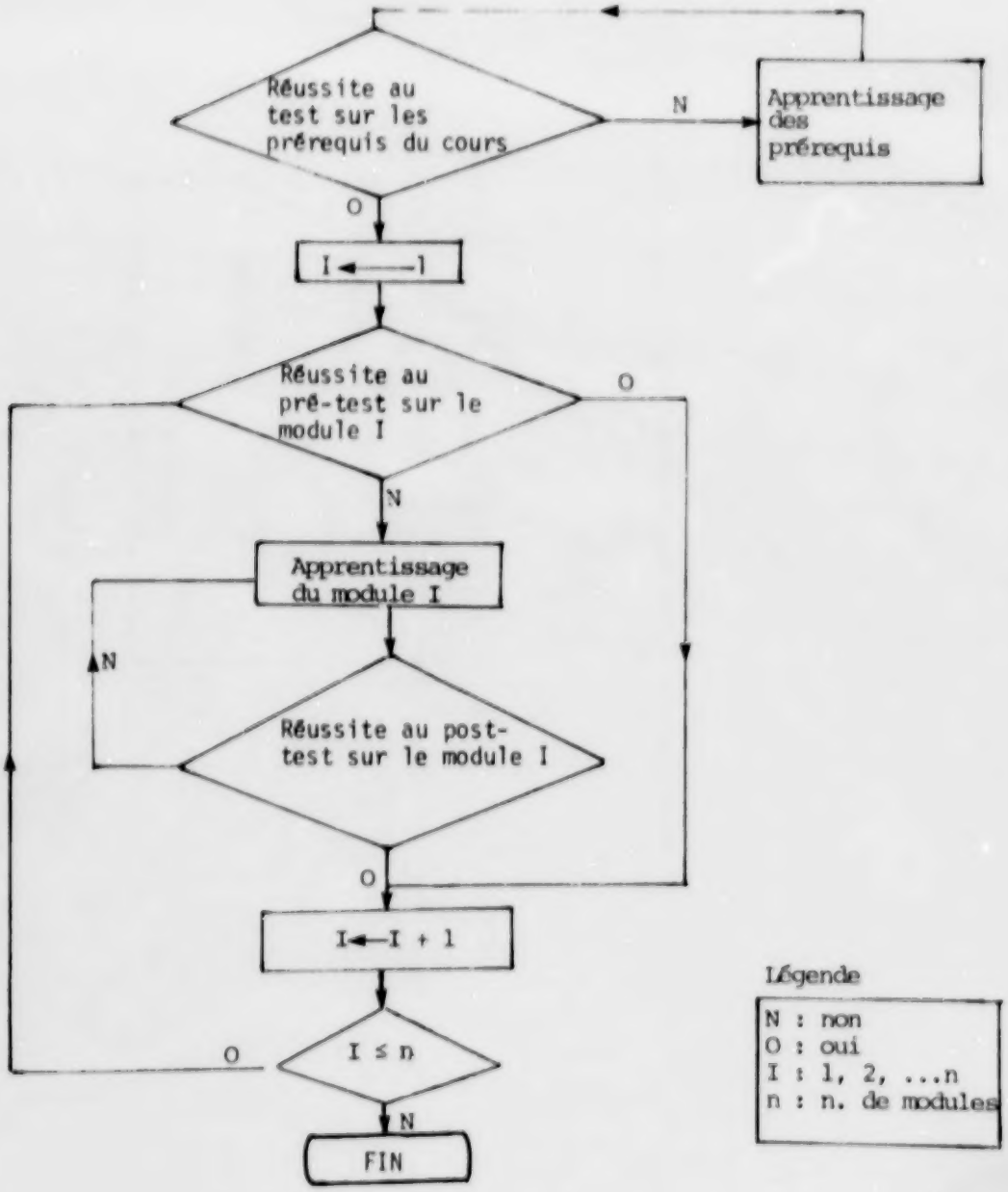
Les étudiants peuvent donc avoir des cheminements différents les uns des autres; ils peuvent avoir des points de départ différents. Ces pré-tests leur évitent de perdre du temps à revoir des parties de matière qu'ils connaissent déjà. Ils peuvent se concentrer sur des apprentissages nouveaux.

Pour un cours donné ou une partie de cours, il peut donc y avoir à la fois évaluation des prérequis et pré-test, car ce sont deux mesures différentes. Examinons certaines possibilités de combinaisons d'évaluation des prérequis et de pré-tests.

Dans l'organigramme suivant, on utilise le test sur les prérequis et le pré-test pour chacune des parties de matière, appelées ici modules, dans un contexte d'évaluation formative.


I ← 1
Réussite au test sur les prérequis du module I
N
Apprentissage des prérequis
O
Réussite au pré-test sur le module I
O
N
Apprentissage du module I
Réussite du post-test sur le module I
N
O
I ← I + 1
I ≤ n
O
N
FIN
Légende
N : Non
O : Oui
I : 1, 2, ...n
n : n. de modules

Il faut cependant dire qu'il ne s'agit que d'un exemple parmi d'autres. Un pré-test ou encore un test sur les prérequis peuvent fort bien ne pas s'appliquer à toutes les sections d'un cours comme on peut les voir s'appliquer à toutes les sections et sous-sections. On peut aussi penser à un seul test sur les prérequis au tout début du cours. L'organigramme suivant représente ce cas dans le cadre d'un enseignement modulaire.



Il est évidemment possible de n'évaluer, pour un cours ou une partie de cours, que les prérequis ou encore de n'utiliser que les pré-tests, selon les besoins. Il y a plusieurs variantes possibles; il s'agit de choisir ou d'inventer celle qui convient.

Notes et références bibliographiques

1. Ces exemples d'objectifs sont inspirés du *Curriculum des études en obstétrique-gynécologie externat-internat*. Montréal, Département d'obstétrique-gynécologie, Faculté de médecine, Université de Montréal, 1978, 211 p..

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Faut-il pénaliser le «choix au hasard»?

France Fontaine

Ce texte s'adresse aux utilisateurs de tests objectifs composés de questions à choix multiple. Une question - ou un item - est à choix multiple lorsqu'elle comprend un énoncé et un choix de réponses. Un test est dit objectif si, à chacune des questions, ne correspond qu'une seule réponse bonne parmi celles qui sont proposées.

Quel utilisateur de ce type de tests n'est pas préoccupé par le problème du choix au hasard? Il se demande s'il doit imposer une pénalité pour le choix d'une mauvaise réponse, appelée aussi leurre. Malheureusement, l'information nécessaire à une prise de décision est peu connue. Nous tenterons de jeter un peu de lumière sur ce sujet.

Pourquoi choisir au hasard?

Dans une question à choix multiple, l'étudiant qui ne connaît pas la réponse peut prendre un risque et répondre correctement en choisissant au hasard l'une des réponses proposées. Le choix d'une réponse au hasard semble encore plus avantageux lorsqu'une mauvaise réponse n'entraîne pas de perte de point. L'étudiant peut alors augmenter sa note ou, au pis aller, conserver la même note en choisissant au hasard chaque fois qu'il ne connaît pas la réponse.

Il est encore plus tentant pour l'étudiant de choisir au hasard lorsqu'il fait face à un test trop difficile ou lorsqu'il n'a pas le temps de répondre à toutes les questions.

En quoi consiste le problème?

On a constaté des tendances très diverses à l'égard du choix au hasard. Les étudiants moins soumis, plus habitués à prendre des décisions ou plus sûrs d'eux-mêmes sont portés à choisir au hasard plus volontiers que les autres, qui sautent plus fréquemment une question dont ils ne connaissent pas la réponse.

Ce sont ces diverses propensions qu'on veut éliminer car elles sont l'expression de variables personnelles qui n'ont généralement pas de lien avec ce qu'on veut mesurer.

De plus, il est difficile de mesurer exactement le degré de connaissance d'un étudiant à partir de ses réponses. L'étudiant a-t-il indiqué la bonne réponse parce qu'il la connaissait? A-t-il choisi un leurre - mauvaise réponse - parce qu'il a choisi au hasard ou parce qu'il croyait, à tort, que c'était la bonne réponse? Quel est le score vrai de l'étudiant au test? Ces interrogations découlent de la possibilité pour l'étudiant de choisir une réponse au hasard.

Le tableau suivant donne diverses raisons qui peuvent amener l'étudiant à choisir une bonne ou une mauvaise réponse, ou encore à ne pas répondre.

Bonne réponse:	- connaissance du sujet - connaissance partielle* - choix au hasard - erreur
Mauvaise réponse:	- choix au hasard - connaissance partielle - ignorance** - erreur
Aucune réponse:	- ignorance - faute de temps

* L'étudiant identifie tout d'abord quelques leurres puis choisit au hasard parmi les réponses qui restent.

**L'étudiant choisit délibérément une mauvaise réponse qu'il croit être la bonne.

Comment corriger l'influence du «choix au hasard»?

Les chercheurs ont proposé diverses solutions à ce problème. Nous en analyserons six, dans les paragraphes qui vont suivre.

Absence de la limite de temps

S'il s'agit d'un test où la limite de temps a pu inciter les étudiants à choisir au hasard pour réussir à répondre à toutes les questions, on peut augmenter la période allouée à l'administration du test, ou diminuer le nombre d'items.

Augmentation du nombre d'options

La probabilité de choisir au hasard la bonne réponse est plus élevée si le nombre de choix proposés est faible: cette probabilité et le nombre d'options sont inversement proportionnels. L'effet du choix au hasard est diminué par l'augmentation du nombre de réponses proposées pour chacune des questions.

Ceci nous incite à proscrire l'usage des items du type «vrai ou faux». Ils n'indiquent pas le degré réel de connaissance d'un étudiant, étant donné la grande part d'erreur qui peut en affecter le résultat. On peut utiliser les questions de ce genre comme moyen d'apprentissage (pour amorcer une discussion, par exemple), mais il faut les éviter dans les tests de certification.

Choix au hasard forcé

Les étudiants reçoivent l'instruction de répondre à toutes les questions. On suscite ainsi chez tous les sujets une égale propension à choisir au hasard. Cependant, cette solution, qui peut sembler parfaite, fait croître les erreurs de mesure. Il semble même que, malgré de semblables instructions, un certain nombre d'étudiants omettent de répondre à certaines questions. De plus, des professeurs refusent cette pratique en invoquant l'effet qu'elle peut avoir sur la perception que se font les étudiants de l'évaluation.

Formule traditionnelle de correction

Il s'agit d'une formule fréquemment utilisée pour estimer le nombre de réponses sues par le répondant à partir du nombre de bonnes réponses qu'il a obtenues au test. Autrement dit, on veut estimer son score vrai au test, son degré réel de connaissance en évaluant l'effet du choix au hasard.

La formule traditionnelle de correction estime le nombre de réponses sues par l'étudiant (X) en retranchant du nombre de bonnes réponses (B) l'estimation du nombre de bonnes réponses obtenues grâce au choix au hasard. On estime le nombre de bonnes réponses dues au choix au hasard à partir du nombre de mauvaises réponses fournies par l'étudiant. La formule traditionnelle de correction est la suivante:

$$X = B - \frac{M}{n-1}$$

où X=estimation du nombre de réponses sues
B=nombre de bonnes réponses
M=nombre de mauvaises réponses
n=nombre de choix par question

Selon cette formule, on estime que l'étudiant qui, dans un examen comprenant 100 questions à 5 choix, répondrait correctement à 70 questions et incorrectement à 16 autres, en laissant 14 questions sans réponse, ne connaîtrait la bonne réponse qu'à 66 questions plutôt qu'à 70. Son score passerait donc de 70 à 66, c'est-à-dire qu'on le pénalise de 4 points pour ses mauvaises réponses. Dans cet exemple, B = 70, M = 16, n = 5 et X = 66.

Examinons cependant les hypothèses sous-jacentes à cette formule qui pénalise le choix au hasard. Elles sont au nombre de deux :

- Si l'étudiant connaît la bonne réponse, il inscrit cette réponse ;
- S'il ne connaît pas la réponse, il ne répond pas ou il choisit aléatoirement parmi les «n» choix possibles (cf. la formule).

La première hypothèse signifie que l'étudiant qui connaît la bonne réponse ne peut se tromper en l'inscrivant. La deuxième hypothèse indique que les diverses options sont également attrayantes pour l'étudiant qui ne connaît pas la bonne réponse : .'étudiant choisit au hasard parmi les «n» options possibles. Ces hypothèses impliquent aussi que les items omis ne sont pas sus (supposant ainsi qu'il n'y a pas de limite de temps) et que toute réponse fausse est le résultat d'un choix au hasard.

La faiblesse de cette formule de correction tient à ses hypothèses peu réalistes et c'est un aspect sur lequel on n'attire pas suffisamment l'attention des utilisateurs. Ainsi, les hypothèses ne tiennent pas compte de la connaissance partielle, c'est-à-dire du fait que l'étudiant peut éliminer certaines options à partir de la connaissance incomplète qu'il a du sujet. Il en découle qu'il est rentable pour l'étudiant qui peut éliminer au moins un leurre de choisir au hasard l'une des options entre lesquelles il hésite car, même lorsque la formule de correction est utilisée, son score espéré est différent de zéro.

Ces hypothèses supposent aussi qu'un étudiant ne peut pas choisir une mauvaise réponse parce qu'il a mal compris l'information donnée au cours. Dans ce cas, il ne s'agit pas de choix aléatoire d'une option, car l'étudiant la choisit en croyant qu'il s'agit de la bonne réponse.

Notons que, comme l'effet du choix au hasard, celui de la formule de correction est inversement proportionnel au nombre de choix de réponses à un item: plus le nombre d'options est grand, plus la pénalité imposée pour une mauvaise réponse est faible.

De plus, on constate que l'utilisation de cette formule influence peu l'ordre des individus forts ou faibles, dans leur groupe respectif, quant aux résultats d'un tel test. Cependant, elle a plus de répercussions chez ceux qui se situent près de la moyenne au test et chez ceux qui réussissent ou échouent selon que la formule de correction est utilisée ou non.

Formule de correction et instruction de «ne pas choisir au hasard»

Pour réduire les diverses tendances à faire des choix au hasard, on compte sur l'effet psychologique de la directive «ne pas choisir au hasard» avertissant

les étudiants qu'ils seront pénalisés pour leurs mauvaises réponses. Or les conclusions des études[1] sur ce sujet semblent toutes indiquer que de telles instructions ne sont pas suivies de la même façon par tous les sujets. Les étudiants au caractère soumis, peu sûrs de leur jugement, peu portés au risque, hésitants à prendre des décisions sont défavorisés, car ils omettent plus d'items que les individus aux caractéristiques opposées; par ailleurs, l'habileté à répondre correctement aux items omis est égale. Et on a quand même des propensions diverses à répondre au hasard selon les individus.

Formule de correction tenant compte de la connaissance partielle

Il existe des formules de correction pour le choix au hasard qui tiennent compte de la connaissance partielle. On a alors un modèle un peu plus réaliste que celui qui sert de base à la formule traditionnelle de correction, car il peut y avoir un choix aléatoire après élimination d'un ou plusieurs leurres selon l'information possédée par l'individu.

Afin de mesurer la connaissance partielle d'un étudiant, on lui donne des instructions différentes sur la façon d'inscrire ses réponses. Par exemple, on peut lui demander d'indiquer le plus petit sous-ensemble de réponses dans lequel il croit que la bonne réponse apparaît; ou encore d'indiquer autant de leurres qu'il peut pour chaque item. L'étudiant est informé qu'il sera pénalisé s'il rejette la bonne réponse, c'est-à-dire s'il l'indique comme étant un leurre. Cette façon de procéder l'assure d'un crédit pour ses connaissances partielles, ce crédit étant fonction du nombre de leurres éliminés.

Etant donné les limites de ce texte, nous n'avons pas l'intention de développer ici les diverses formules de correction pour le choix au hasard qui tiennent compte de la connaissance partielle. Disons seulement qu'on trouve aussi dans ce cas que ce sont les étudiants prudents, moins sûrs d'eux, qui omettent le plus d'items, donc qui sont pénalisés[2].

Mais alors, que faire?

Etant donné les diverses restrictions à observer quant à l'utilisation des formules de correction, nous recommandons de ne pas pénaliser le choix au hasard.

Le problème du choix au hasard ne tient pas seulement au calcul des probabilités. Lorsqu'on utilise des formules de correction pour le choix au hasard, il faut avoir en tête les hypothèses sous-jacentes à ces formules, hypothèses qui ne peuvent, même dans les cas les plus sophistiqués, tenir compte de toute la réalité. De plus, ces formules associées à diverses instructions ne peuvent corriger complètement les effets des diverses tendances à choisir au hasard. Ce sont toujours ceux qui ont moins confiance en eux qui sont pénalisés. C'est pourquoi plusieurs auteurs proposent l'abandon de ces formules qui cherchent à diminuer l'effet du choix au hasard.

En résumé, nous sommes aussi de l'avis de ne pas pénaliser le choix au hasard et nous incitons à la prudence ceux qui désirent tout de même utiliser ces formules.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 3 No 4, Déc. 1978, pp. 12-15.

1.Rouley, G.L., «Which examines are most favoured by use of multiple choice tests» in *Journal of Educational Measurement*, vol. II, no 1, 1974, pp. 15-23.

Sherriffs, A.C. and Boomer, D.S., «Who is penalized by the penalty for guessing?» in *Journal of Educational Psychology*, vol. IV, 1954, pp. 81-90.

Slakter, M.J., «The penalty for not guessing» in *Journal of Educational Measurement*, vol. 5, 1968, pp. 141-144.

Votaw, D.F., «The effect of «do not guess» directions on the validity of true-false and multiple-choice tests» in *Journal of Educational Psychology*, vol. XXVII, 1936, pp. 698-702.

2.Slakter, M.J., «The penalty for not guessing» in *Journal of Educational Measurement, vol. 5, 1968, pp. 141-144.*

Section 4

La notation

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Les systèmes de notation

Huguette Bernard

Il arrive assez fréquemment, dans la vie d'une institution d'enseignement, qu'on se demande si on doit ou non changer de système de notation. Les nombreuses études sur le sujet indiquent à elles seules que le choix d'un système de notation est un des thèmes favoris en évaluation.

La première partie de ce texte sera donc consacrée aux divers systèmes de notation utilisés par les universités canadiennes et américaines. Nous verrons, par exemple, les systèmes de notation dits traditionnels, tels que le système numérique et le système littéral et les systèmes de notation dits non traditionnels, tels que le «pass/fail» (succès/échec) et ses variantes. Pour chacun d'eux, nous ferons une analyse des avantages et des inconvénients.

En second lieu, nous verrons pourquoi les systèmes de notation ont été inventés. Nous étudierons si les systèmes de notation satisfont vraiment les besoins pour lesquels ils existent. C'est à partir des données fournies par une revue de la recherche en ce domaine que nous ferons cette critique.

Nous ferons ensuite le point sur le problème de la notation à partir de ces études sur les divers systèmes de notation et sur l'utilité des notes. Nous essayerons de donner un éclairage nouveau à ce problème du choix d'un système de notation.

D'abord, quelques définitions

La note: C'est le symbole numérique (50%, 70%, 90% ...) ou littéral (A, B, C, ... CR, P, NC, ..) signifiant le résultat obtenu par un étudiant pour un cours, ou un ensemble de cours pendant une période donnée.

Le rapport: C'est l'information que l'on veut transmettre aux étudiants, aux employeurs, aux bureaux d'admission d'une école supérieure, sur les résultats d'un étudiant aux cours suivis.

La mesure: C'est la collecte et la classification systématique de données quantifiables sur un objet ou une classe d'objets; ici l'objet peut être l'apprentissage d'un étudiant.

L'évaluation: C'est le jugement que l'on porte face aux résultats des étudiants.

LES DIFFÉRENTS SYSTÈMES DE NOTATION

Nous distinguerons ici les systèmes dits traditionnels, c'est-à-dire ceux qui sont véhiculés depuis toujours dans notre enseignement: le système numérique par pourcentage et le système littéral; les systèmes dits non traditionnels, tels que le «pass-fail» et ses variantes.

Les systèmes dits traditionnels

Le système numérique par pourcentage et le système littéral ont évolué avec le temps. On les trouve en général hermétiques, énigmatiques et sans signification.

Le système numérique par pourcentage

Voyons un peu quels sont les codes ou symboles par lesquels on exprime les résultats académiques des étudiants. Certaines institutions utilisent le système numérique par pourcentage. Sur le rapport de notes (le bulletin) des étudiants, à chaque cours correspond un pourcentage. Habituellement, l'étudiant qualifié de brillant obtiendra un pourcentage supérieur à 80%, l'étudiant moyen se situera entre 60% et 80% et l'étudiant faible échouera avec un pourcentage inférieur à 60% ou 50%, selon les politiques des facultés et des départements. Dans tous les cas, le pourcentage attribué à l'étudiant peut provenir des résultats d'un seul examen, de plusieurs examens, de travaux et ainsi de suite. C'est alors la mesure qui est à la base de la note. Parfois la note est la transcription d'un jugement porté par le professeur sur l'étudiant; ce jugement peut viser ses capacités ou son intérêt, ou encore sa ponctualité et le reste.

Le système littéral

D'autres institutions emploient le système littéral simple, c'est-à-dire celui par lequel on indique les résultats des étudiants au moyen des lettres de l'alphabet A, B, C, D, E, F. Évidemment, ce système permet moins de discrimination entre les étudiants: il y a moins de notes possibles. Pour certains, c'est un désavantage, pour d'autres, cela traduit une réalité, puisqu'on ne saurait évaluer à 1% près les capacités d'un étudiant. De plus, dans ce système, il est difficile de faire une moyenne des notes; cependant, le problème est habituellement contourné par la conversion des lettres en pourcentages, puis par leur reconversion en lettres.

Ce système, comme tous les autres du genre qui suivront, n'échappe pas au manque de précision de la note. La lettre, tout comme le pourcentage, n'indique pas si elle provient d'un ou plusieurs examens, d'un travail, d'une combinaison de travaux et d'examens, d'une présence obligatoire au cours et le reste. Toutes les variétés de données quantifiables sont possibles. Ces systèmes n'indiquent pas les apprentissages qui sont maîtrisés et ceux qui ne le sont pas.

Ces deux systèmes (numérique et littéral) semblent posséder les mêmes inconvénients. Par exemple, on a dû, dans les deux cas, discuter de la limite de passage ou d'échec. Doit-elle se situer à 60% ou à 50%? A D, E ou F? Tout dépend des exigences de chaque département.

La combinaison des systèmes de notation

On a tenté de combiner ces deux systèmes de notation afin d'obtenir les avantages des systèmes numérique et littéral. On a pensé qu'en regroupant les deux systèmes, on atteindrait une plus grande objectivité et même une meilleure compréhension, et de la part des professeurs qui distribuent les notes, et de la part des étudiants qui les reçoivent. On a tenté par exemple d'accoler un pourcentage à une lettre; de plus, en combinant ces deux systèmes, on a voulu qualifier le résultat: excellent, bien, passable, ...

Les tableaux qui suivent illustrent quelques exemples des systèmes combinés de notation traditionnels. En premier lieu, on présente un tableau extrêmement complexe de la combinaison du système littéral et du système numérique par pourcentage. Dans le second exemple, nous verrons une simplification du pairage lettre - pourcentage, auquel on aura ajouté une mention ayant pour but de qualifier la note. Enfin dans un troisième temps, nous illustrerons le pairage lettres-mentions.

Pairage lettres - pourcentages

LETTRES	%	ECART
A^+	90-100	11
A^-	85-89	5
A^-	80-84	5
B^+	77-79	3
B^-	73-76	4
B^-	70-72	3
C^+	67-69	3
C^-	63-66	4
C^-	60-62	3
D^+	57-69	3
D^-	53-56	4
D^-	50-52	3
F^+	42-49	8
F^-	35-41	7
F^-	0-34	35

Tableau 1
Équivalence lettres-pourcentages

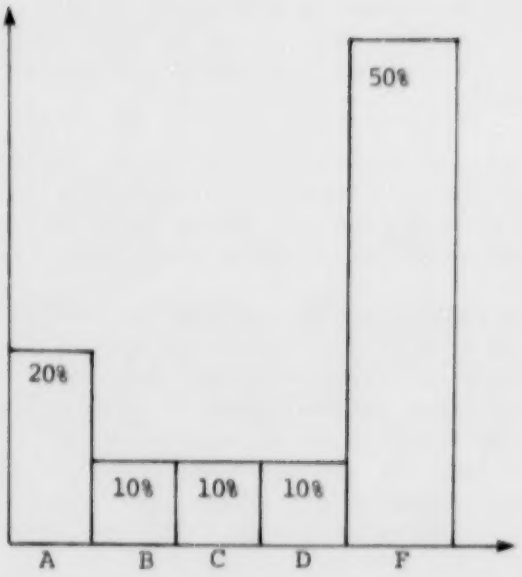


Figure 1
Représentation graphique des écarts de pourcentages pour chacune des lettres regroupées sans signe.

Il est important de noter ici les écarts variables d'une lettre à une autre. L'ensemble des lettres A totalise 20 points; l'ensemble de lettres B, C, ou D totalise chacun 10 points; enfin l'ensemble des lettres F totalise à lui seul 50 points. L'histogramme (figure 1) illustre davantage la répartition des notes de ce système utilisant des lettres en correspondance avec des pourcentages.

Mais un des inconvénients de ce système est évidemment qu'il faut toujours consulter la table de correspondance pour le comprendre. Ni le professeur, ni les étudiants, et encore moins les administrateurs, ne peuvent dire exactement la signification de l'une ou l'autre des lettres complétées de son signe. Par exemple, à quoi correspond exactement la lettre C? Dans la table, C se situe entre 63% et 66%; sans la table, on peut supposer que C équivaut à l'écart de 60% à 70%. Un second inconvénient est l'écart entre chaque lettre. Par exemple, comment attribuer un C+ par comparaison avec un B- en sachant que la lettre B est plus prestigieuse que la lettre C, quel que soit son signe.

Pairage lettres - pourcentages - mentions

A	76-100%	" First class honors "
B	66-75	" Second class honors "
C	60-65	" Third class honors "
D	50-59	" Passing "
F	50 et -	" Failure "

Tableau 2
Équivalence lettres, pourcentages et mentions.

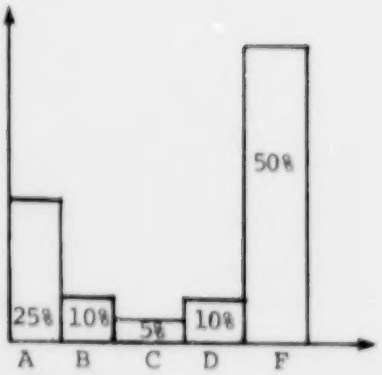


Figure 2
Représentation graphique des écarts de pourcentages pour chacune des lettres.

Ici, on remarque (figure 2) que les écarts entre chacune des lettres diffèrent de l'exemple précédent (figure 1). Il y a plus de pourcentages représentés par la classe A, et moins par la classe C. C'est donc dire que l'étudiant qui a obtenu un A avec 76% obtient somme toute la même note que celui qui a réussi avec 100%. Les deux obtiennent la mention «first class honors». De plus, la classe C ne recouvre que cinq notes possibles.

Avec cette variante des systèmes traditionnels, on essaie de qualifier la compétence de l'étudiant en même temps qu'on la quantifie. Ce système concilie à la fois trois écoles de pensée: l'école en faveur des pourcentages, l'école en faveur des lettres et l'école en faveur d'une appréciation qualitative plus globale.

Pairage lettres-mentions

Les trois exemples qui vont suivre se ressemblent; seul le vocabulaire change. On remarquera cette fois que les pourcentages ont été complètement abandonnés. De plus, on notera l'arrivée de la lettre E pour les deux premiers systèmes et l'abandon des lettres D et F, qui sont remplacées par d'autres lettres n'ayant absolument pas la même interprétation dans le troisième exemple.

A	excellent
B	bon
C	acceptable
D	faible
E	insuffisant
F	reprise

Tableau 3.1

A	excellent
B	bien
C	moyen
D	passable
E	échec
F	reprise

Tableau 3.2

A	" high competence "
B	" good competence "
C	" moderate competence "
CR ou P	" satisfactory achievement of course objective "
NC	" no credit ", " competence not demonstrated "

Tableau 3.3

Équivalence lettres - mentions

Dans les systèmes qui ont précédé, D était toujours la limite de passage. Ici, on introduit dans le système français la lettre E, qui peut être interprétée comme une note contestable ou sujette à révision, alors que F implique automatiquement la reprise.

Enfin, dans le dernier système, les lettres A, B et C ne varient pas. Elles correspondent aux mêmes lettres que précédemment. Les lettres CR ou P indiquent que l'étudiant a satisfait aux exigences du cours et les lettres NC indiquent que les crédits sont refusés, l'étudiant n'ayant pas satisfait aux exigences du cours. Ce qui est intéressant, dans ce dernier exemple, c'est la nouveauté du langage. Par exemple, on dit de l'étudiant qu'il a une bonne compétence en la matière ou qu'il a répondu de façon satisfaisante aux objectifs du cours.

Si on observe l'évolution des systèmes de notation traditionnels, on remarque que, de plus en plus, on a essayé de rendre moins abstraits les pourcentages ou les lettres et qu'on a essayé de juxtaposer une information aux notes: on pose un jugement. On peut dire d'un étudiant qu'il est «hautement compétent», «compétent» ou que sa compétence n'a pas été démontrée dans tel ou tel domaine. C'est d'ailleurs le souvenir qu'on garde d'un étudiant lorsqu'il a quitté l'université.

Cependant, on explicite rarement les critères d'évaluation qui permettent d'indiquer une note, qu'il s'agisse d'un A, d'un 80% ou de la mention «compétent», au dossier de l'étudiant. Peut-être le professeur les connaît-il? Il est moins sûr que l'étudiant ou toute autre personne intéressée les connaissent. Quelle capacité l'étudiant démontre-t-il réellement lorsqu'il obtient un 80% à un examen de mathématiques ou d'histoire?

Les systèmes dits non traditionnels

Dans les systèmes de notation non traditionnels, nous parlerons explicitement du système «pass-fail» (succès-échec) et de ses variantes. Il est à noter que le système «pass-fail», tel que vécu dans les expériences mentionnées dans ce texte, n'apparaît pas comme un système d'évaluation basé sur des objectifs bien précis et possédant des critères nettement établis. Ce serait plutôt un jugement global qu'on donnerait à la fin d'un cours.

Le «pass-fail»

Le système de notation «pass-fail» (succès-échec) consiste à accorder la note «pass» à l'étudiant lorsqu'il a satisfait aux exigences du cours et la note «fail» dans le cas contraire.

Il n'y a que ces deux mentions qui peuvent être inscrites au dossier de l'étudiant; dans un tel système, la compétition n'existe plus. Quel que soit le rendement des étudiants (sauf s'ils échouent), ils obtiennent la même note. À l'origine, le système «pass-fail» a été mis en oeuvre pour permettre aux étudiants de s'inscrire à des cours en dehors de leur spécialité. Aussi les premières universités à mettre de l'avant cette option conservaient leur propre système de notation pour leurs cours réguliers et adoptaient le «pass-fail» pour les cours optionnels. La raison d'une telle pratique était d'encourager l'étudiant à diversifier son champ d'études puisqu'un succès moindre dans un domaine autre que celui de sa spécialité ne risquait pas d'abaisser sa moyenne. Certaines universités restreignent cette notation à un seul cours par étudiant; d'autres ont appliqué ce système à tous les cours.

Évidemment, l'application partielle ou totale d'un tel système dépend de ce qu'on en attend. Pour certains, le système «pass-fail» n'a que des avantages, comme ceux-ci par exemple:

- encourager les étudiants à explorer d'autres domaines;
- éliminer la compétition et réduire l'anxiété;
- centrer les efforts sur l'apprentissage plutôt que sur les notes;
- donner à l'étudiant la maîtrise de son emploi du temps;
- éliminer de la fonction d'enseignant le rôle de juge.

Certains étudiants sont plus curieux que d'autres et désirent étendre leur champ de connaissance; le système «pass-fail» les encourage à le faire. De plus, l'étudiant n'a pas à craindre de voir baisser ses notes puisque, s'il satisfait aux exigences du cours, il obtiendra la note de passage. De cette façon, son anxiété possible face au cours s'en trouve diminuée[1]. Donc, l'étudiant arrivera peut-être à centrer son effort sur l'apprentissage lui-même plutôt que sur les notes. De plus, l'étudiant se voit responsable de son

emploi du temps. Il peut décider de consacrer moins d'heures à certains cours d'une importance moindre pour lui et de se concentrer davantage sur une matière dont l'importance lui est primordiale.

Ceux qui ont généralisé à tous les cours ce système de notation y voyaient des avantages certains. Pour eux, la liberté de l'étudiant était prioritaire, de même que son bien-être. Les instigateurs de ce système avaient décidé de changer la mentalité des professeurs et des étudiants. Ils éliminaient de la fonction de professeur le rôle de juge et rendaient à l'école sa fonction première, celle d'apprendre à apprendre et non pas de donner des notes.

Mais le système ne semble pas parfait; ses détracteurs lui prêtent les inconvénients suivants:

- Il élimine tout esprit de compétition chez les étudiants;
- Il affaiblit la motivation de l'étudiant et son apprentissage est par le fait même diminué;
- les cotes attribuées ne peuvent rentrer dans le calcul de la moyenne et, si le nombre de cours optionnels est considérable, la moyenne n'est plus significative;
- l'étudiant «doué» fait le minimum d'effort pour obtenir la note de passage afin de ménager ses énergies pour les cours ayant conservé le système traditionnel de notation;
- Il rend difficile l'accès à d'autres établissements scolaires ou professionnels.

La compétition est-elle nécessaire dans le domaine scolaire? Pour certains étudiants, les notes traditionnelles les poussent à travailler; pour d'autres, c'est le contraire. Des études ont montré que les étudiants qui réussissent bien en classe ont une haute opinion d'eux-mêmes face à leur réussite, tandis que les étudiants moins talentueux se sentent diminués par leurs notes faibles et souvent préfèrent abandonner leurs études.

Quelle est la place de la motivation dans un système «pass-fail»? Lors d'une étude[2], il a été démontré que, pour le tiers des étudiants interrogés, leur motivation avait subi une baisse dans un tel système de notation, alors que seulement 12% ont senti leur motivation augmenter. De même, le quart de ces étudiants ont eu l'impression d'avoir moins appris dans le système «pass-fail», contre 13% qui ont déclaré avoir appris davantage. Pour la majorité des étudiants interrogés, rien n'a changé: ni leur motivation, ni leur façon d'apprendre.

Cette même étude indique également que le but premier de ce système de notation n'a pas été réalisé: c'est-à-dire que le «pass-fail», au lieu d'encourager les étudiants à explorer des nouveaux secteurs de la connaissance, les a tout simplement amenés à vouloir avant tout protéger leur moyenne en essayant d'avoir une bonne note pour les autres cours.

Une seconde étude[3] sur ce sujet a établi que le système «pass-fail» utilisé pour tous les cours pendant une année provoquait une diminution de la performance académique. Ainsi les étudiants qui sont retournés au système traditionnel après une année de «pass-fail» ont obtenu des notes plus faibles que les étudiants qui sont restés dans le système traditionnel.

Une étude de Bishop[4] indique que 60% des étudiants disent avoir travaillé aussi fort, sinon plus fort, dans leurs cours utilisant le «pass-fail» que dans les autres cours, tandis que 40% avouent qu'ils ont travaillé moins fort dans ces cours. Cependant seulement 5% des étudiants disent avoir moins retiré de leurs cours notés «pass-fail» que de leurs autres cours, alors que près de 60% indiquent en avoir autant retiré, et que près de 30% affirment en avoir retiré davantage. Les réactions des professeurs et des étudiants au système «pass-fail» étaient positives.

Enfin, d'autres études[5] ont révélé que les étudiants utilisent le système «pass-fail» dans le seul but de mieux distribuer leur temps et leur effort, de façon à se concentrer davantage sur les cours comportant un système de notation traditionnel. Il apparaît donc que de changer le système de notation pour une partie des cours, en laissant le reste inchangé, peut produire des effets indésirables, vu la pression qui incite les étudiants à obtenir de bonnes notes pour ces autres cours.

Et finalement, il a été démontré[6] qu'un candidat a moins de chances d'être choisi pour un poste professionnel si son rapport contient plus de 10% de notes non traditionnelles. Dans cette même optique, on a prouvé qu'un candidat ayant plus de 10% de notes non traditionnelles dans son rapport voit ses chances diminuer lorsqu'il tente de s'inscrire dans une institution supérieure d'éducation; ses chances de financement au moyen d'une bourse ou d'un prêt sont pareillement réduites.

Il est à noter que toutes les recherches mentionnées précédemment portent sur le système «pass-fail», qui accorde la note de réussite ou d'échec, de façon globale, pour un cours. Le «pass-fail» expliqué ici ne résulte pas de la mesure d'un apprentissage évalué à plusieurs reprises selon des objectifs bien précis et des critères nettement établis.

Les variantes du système «pass-fail»

Les adeptes du «pass-fail» ont tenté d'affiner le système. C'est ainsi qu'on a vu naître le

Honors - Pass - Fail et le Distinction - High pass- Pass

Dans ce dernier cas, on élimine le côté négatif de l'évaluation, c'est-à-dire que l'échec n'est pas inscrit au dossier de l'étudiant, ce qui lui permet une reprise. C'est une pratique assez courante même dans le système de notation traditionnel: on inscrit au dossier de l'étudiant la mention «incomplet», lorsqu'il a échoué à un examen ou lorsqu'il tarde à remettre un travail. De cette façon, on lui donne la chance de se reprendre. De plus, les modifi-

cations du système «pass-fail» semblent justifier le besoin de compétition chez les étudiants ou de discrimination de nos institutions puisqu'on hiérarchise les nouveaux systèmes. On en revient presque à un système littéral simplifié.

Dans cette première section, nous avons vu les systèmes de notation dits traditionnels et les systèmes de notation dits non traditionnels. Les seconds, à ce qu'il semble, ont les inconvénients des premiers, c'est-à-dire qu'ils ne renseignent pas sur les apprentissages précis maîtrisés ou non par les étudiants, car le jugement est global. Les diverses études nous laissent insatisfaits; nous ne sommes pas convaincus de la supériorité d'un système par rapport aux autres. Afin d'éclairer ce problème, passons à la seconde section, qui précisera le pourquoi de la notation.

L'UTILITÉ DES NOTES

Rappelons d'abord les buts des systèmes de notation, afin d'établir si la notation satisfait bien les besoins pour lesquels elle existe.

Les notes servent à des fins administratives

Les notes indiquent si l'étudiant a réussi ou échoué, s'il est promu ou s'il doit reprendre son cours, ou encore si l'on doit lui décerner un diplôme. Les notes sont utiles aussi pour le transfert d'un étudiant d'une institution à une autre; elles servent à son admission dans un établissement scolaire. Enfin elles guident les employeurs pour l'engagement de futurs employés.

Les notes jouent un rôle d'orientation

Les notes aident à identifier les domaines où l'étudiant réussit le mieux et ceux où il réussit le moins bien. Ainsi elles facilitent son orientation. De plus, elles peuvent déterminer le nombre de cours auxquels l'étudiant peut choisir de s'inscrire.

Les notes assument un rôle d'information

Par le rapport, les notes indiquent à l'étudiant et aux parents, ou autres personnes intéressées s'il y a lieu, les progrès, succès ou échecs de l'étudiant.

Les notes jouent un rôle de motivation et d'auto-discipline

Les notes stimulent l'étudiant et le portent à fournir de plus grands efforts dans son activité d'apprentissage.

A quel point ces rôles sont-ils remplis?

Examinons, à la lumière de diverses recherches[7], si la notation répond adéquatement aux rôles qu'on lui a attribués.

Là où il est prouvé que les notes ne prédisent pas le succès

Les notes peuvent-elles prédire le succès professionnel? Des études ont démontré que, dans des domaines professionnels bien précis, les notes n'avaient rien à voir avec le succès professionnel. Par exemple, en médecine, une très faible relation a été constatée entre les notes et le succès professionnel. Par contre, en physique, aucune relation n'a pu être établie entre les notes et vingt-quatre énoncés de critères de performance d'un bon physicien[8]. Même chez les ingénieurs, il existe très peu de relation entre les notes et la prédiction du succès professionnel. De plus, cette étude a démontré qu'il n'existait aucune relation entre les notes et le salaire de l'ingénieur[9].

Les notes sont-elles vraiment utiles aux employeurs? C'est plutôt la performance pendant les premières années d'emploi qui a de la valeur prédictive. *Les notes ne peuvent que faiblement prédire le succès scolaire et plus faiblement encore le succès professionnel*. Dans la région de Chicago[10], les employeurs d'entreprises privées et du gouvernement ont affirmé que les notes venaient en second lieu, après l'expérience, dans le choix de leurs candidats.

Là où il est prouvé que les notes ont peu de signification

Des études ont démontré combien la signification des notes diffère d'un professeur à un autre. De plus, il a été prouvé que la signification d'une note ne diffère pas seulement d'un établissement scolaire à un autre, d'un département à un autre et d'un professeur à un autre, mais aussi chez un même professeur, lorsqu'il attribue ses notes à des moments différents pour un même travail.

Par exemple[11], cent professeurs ont corrigé la même copie en anglais, langue maternelle; les résultats ont révélé 39 points d'écart. On a tenté de justifier cet écart en parlant de matière subjective. On peut penser par exemple que dans cette matière, certains professeurs se refusent à attribuer des notes plus élevées que 75 ou 80%, ce qui n'est pas le cas de tous.

On a alors expérimenté la chose avec cent professeurs de géométrie[12]; les résultats obtenus indiquaient un écart de 45 points. Comment expliquer la chose? Plus tard, en 1930, 1952 et 1967, d'autres expériences du même genre ont abouti à des conclusions similaires.

Ces résultats peuvent s'expliquer par le fait que les notes sont souvent une combinaison de résultats scolaires (tels les tests ou les examens, qui eux-mêmes sont d'une fidélité et d'une validité variable) et de jugements du professeur (tels l'effort de l'étudiant, sa ponctualité, son comportement, la présentation et la correction grammaticale de son travail...). Très souvent, le jugement du professeur sur des critères autres que la performance de l'étudiant l'emporte, surtout aux premiers niveaux d'enseignement.

On constate donc que les professeurs notent les étudiants de façon variable et souvent de façon non mesurable. Par exemple, certains professeurs avouent qu'ils font échouer l'étudiant qui ne semble pas intéressé, ne fait pas d'efforts suffisants (même s'il est un des plus brillants), n'est pas attentif pendant les cours ou est trop souvent absent. Par ailleurs, on a constaté que les critères de réussite clairement identifiés étaient rares et les critères d'échec arbitraires. Ou encore, les critères ne correspondent pas aux apprentissages: par exemple, dans un stage, on évaluera la ponctualité, les connaissances théoriques, le port de l'uniforme, etc.

Il s'ensuit donc qu'obtenir une note élevée ou la mention «pass» («succès») ne signifie pas nécessairement qu'on a acquis des compétences. Cependant, un échec est toujours un échec et, que les critères soient arbitraires ou non, il reste inscrit en permanence dans le dossier de l'étudiant.

Là où il est prouvé que les notes informent peu

En plus du fait que le peu de signification des notes diminue l'information qu'on peut en tirer, on ne trouve habituellement dans un rapport étudiant que des moyennes. La moyenne réduit l'information sur la performance d'un étudiant. On a même, à un certain moment, été jusqu'à cumuler les notes de l'étudiant en français, en mathématiques, en histoire, etc. Comment penser que cette moyenne nous donnera le profil exact de l'étudiant? Dans un cours comportant des apprentissages pratiques (laboratoire, stage, etc.) et des apprentissages théoriques, on fait une moyenne et on perd ainsi une information précieuse.

Là où il est prouvé que les notes ne stimulent pas les étudiants

Arrêtons-nous d'abord au problème de la motivation pédagogique. Qu'est-ce donc que la motivation pédagogique? C'est la présentation d'un contenu d'étude ou d'un modèle de conduite apte à solliciter l'intérêt et à soutenir un engagement volontaire de l'étudiant. Pour qu'il y ait motivation, il faut une tension mais, au-delà d'un certain seuil, l'essentiel de l'énergie est utilisé pour combattre l'anxiété et, par voie de conséquence, limite la volonté d'apprendre[13].

En ce sens, une recherche faite il y a 10 ans indique que les étudiants qui obtiennent de faibles notes continuent à échouer. Chez certains, l'anxiété augmente la performance, mais chez d'autres, elle engendre l'échec. De plus, on stipule que l'étudiant qui rate un examen devrait revoir sa matière et reprendre l'examen, ne serait-ce que pour témoigner de ses progrès[14].

Des recherches plus récentes ont mis en lumière les effets différents de notes faibles, selon les individus. On rapporte que les étudiants qui reçoivent les notes D et F, comparativement à ceux qui obtiennent des notes plus élevées, ont plutôt tendance à abandonner leurs cours. Mais par contre, les étudiants qui persistent malgré leurs faibles notes réussissent mieux aux examens ultérieurs. On a remarqué de plus que les étudiants qui obtiennent des notes fortes au premier examen ont une meilleure performance au second[15]. Des études ont prouvé que, contrairement à l'opinion populaire, les notes stimulent très

faiblement la plupart des étudiants. Habituellement, les étudiants les plus stimulés par leurs notes sont ceux qui réussissent le mieux. On rapporte que la tension augmente les résultats des étudiants qui sont assurés de bonnes notes, alors qu'elle diminue les résultats des étudiants qui obtiennent normalement des notes moyennes[16]. Ainsi les notes peuvent être le facteur déterminant de l'estimation que l'étudiant fait de lui-même face à sa réussite.

D'autre part, certaines études ont établi que les notes ne sont pas essentielles à la motivation des étudiants. Une recherche indique même le contraire. Les résultats de cette recherche tendent à établir que les notes peuvent nuire au développement intellectuel et aux habiletés personnelles des étudiants[17]. Enfin d'autres recherches ont révélé que les notes peuvent avoir des effets indésirables. Par exemple, elles produisent un effet de conformité chez l'étudiant[18] (celui-ci fait ce que le professeur demande); de plus, elles réduisent l'interaction professeur-étudiants et, finalement, elles poussent les étudiants à tricher afin de s'assurer la moyenne de passage[19].

Revenons à la question posée au début: la notation remplit-elle bien les rô qui lui sont assignés? Nous pouvons, à la lumière de ces recherches, douter de la façon dont les divers systèmes de notation s'acquittent de leurs rôles. En effet, il a été démontré que la note ne prédit pas le succès; de plus, la note n'a plus aucune signification si elle n'est pas basée sur des critères. Elle est habituellement subjective et laisse place à de multiples interprétations.

Conclusion

Plusieurs institutions d'enseignement se sont demandé quel système de notation choisir parmi tous ceux qui leur étaient proposés; elles ont parfois remplacé leur système de notation par un autre. Étudier la question sous cet angle ne nous semble pas devoir amener de solution. Cette façon de poser le problème uniquement en comparant les divers systèmes de notation nous apparaît illusoire: les systèmes semblent se valoir les uns, les autres... C'est que le problème ne réside pas essentiellement dans le choix d'un système de notation.

Il vaut mieux concentrer ses efforts à définir ce pourquoi on veut évaluer, choisir le modèle d'évaluation qui convient, définir les objectifs d'apprentissage, préciser les critères d'évaluation correspondant à ces apprentissages, puis se demander, après toutes ces démarches, quel est le système de notation qui correspondrait le mieux aux buts de l'évaluation, aux modalités choisies.

Les notes semblent trop fréquemment vidées de leur sens pour que l'on puisse croire améliorer les choses de façon tangible en changeant simplement de système de notation. C'est plutôt toute une démarche de révision de l'évaluation qu'il faut entreprendre, la notation n'en étant que la dernière étape.

Notes et références bibliographiques

1 Bains, P.T., Hales, L.W., Rand, L.P., «Does Pass-Fail Encourage Exploration?» in *College and University* 47, Fall 1971, pp. 17-18.

2 Id., ibid.,

3 Gold, R.M., Reilly, A., Silberman, R. Lehr, R., «Academic Achievement Declined Under Pass-Fail Grading» in *Journal of Experimental Education* 39, Spring 1971, pp. 17-21.

4 Bishop, L.K., *Individualizing Educational Systems*. New-York, Harper and Row, 1971, pp. 192-199.

5 Pedrini, B.C. and Pedrini, D.T., *Pass-Fail Grading: Summary and Tentative Conclusions*. Omaha, Nebraska, University of Nebraska at Omaha, 1972.

6 Thomas, J.E. et al., «Effects of Non-Traditional Grades on Hiring Practices of Schools Systems» in *Journal of Educational Measurement* 11, Fall 1974, pp. 213-217.

7 La majeure partie de ces informations sont contenues dans: Evans, Francis B., «What Research Says about Grading» in *Degrading the Grading Myths: A Primer of Alternatives to Grades and Marks*. Washington, Association for Supervision and Curriculum Development, 1976, 151 p.

8 Prince, P.B., Taylor, C.W., Richards, Jr. J.M., Jacobsen, T.L., *Performance Measures of Physicians, Final report submitted to the U.S. Office of Education*. Washington, D.C., 1963.

9 Martin, R.A., Pacheres, J., «Good Scholars Not Always the Best» in *Business week*, February, 1962, pp. 77-78.

10 Bailey, R.L., *AACRAO Subcommitee on Non-Traditional Grading Patterns*. Park Forest South, Illinois, Governors State University, 1972.

11 Starch, D. and Elliot, E., «Reliability of the Grading of High School Work in English» in *School Review* 20, 1912, pp. 442-457.

12 Starch, D., Elliot, E., «Reliability of the Grading Work in Mathematics» in *School Review* 21, 1913, pp. 254-295.

13 Gagnon, M., *La motivation*. Montréal, Service pédagogique, Université de Montréal, 1976.

14 Chansky, N.M., «The X-Rays of the School Mark» in *Educational Forum*, 1962.

15 Thayer, R.E., «Do Low Grades Cause College Students To Give Up?» in *Journal of Experimental Education* 41, Spring 1973, pp. 71-73.

16 Phillips, B., «Sex Social Class and Anxiety as Sources of Variation in School Anxiety» in *Journal of Educational Psychology* 53, 1962, pp. 316-322.

17 Chamberlin, C., Chamberlin, E.S., Drought, N.E. and Scott, W.E., «Did they succed in College?» in *Adventure in American Education*, Vol. 4. New-York, Harper and Rou, 1942.

18 Bostrom, R.N., Vlandis, J.W., Rosenbaum, M.E., «Grades as Reinforcing Contingencies and Attitude Change» in *Journal of Educational Psychology* 52, 1961, pp. 112-115.

19 Bowers, W., *Student Dishonesty and Its Control in College*. New-York, Bureau of Applied Behavioral Science, 1964.

Abolir les notes : pourquoi pas?

France Fontaine

Nous allons présenter divers problèmes suscités par la notation et son usage au niveau universitaire, pour ensuite nous demander comment il pourrait être possible d'envisager l'abolition des notes. Mais, auparavant, nous définissons ce que nous entendons par la notation.

Qu'est-ce que la notation?

La notation est l'utilisation de symboles indiquant le rendement académique des étudiants. Nous savons tous que le code en usage dans notre institution est celui des chiffres de 0 à 100 et que nous retrouvons, dans d'autres institutions, des systèmes de notation utilisant des lettres comme A, B, C, ... ou selon la formule «succès/échec».

La note versée au dossier scolaire de l'étudiant représente, soit une partie de l'évaluation (ou du jugement de valeur) du rendement d'un étudiant, soit, plus directement, la mesure elle-même faite à partir d'un instrument de mesure.

Problèmes posés par la notation

La notation soulève cependant divers problèmes. L'un des plus graves, comme nous l'entendons souvent dire, c'est que les étudiants travaillent pour des notes. Cela peut donner lieu à un «bourrage de crâne» à la veille des examens ou amener un étudiant à se féliciter d'avoir réussi un examen, par chance, puisqu'il ne portait que sur les parties de la matière que l'étudiant connaissait. Ces comportements ne sont peut être pas ceux que nous attendions de nos étudiants, mais qui peut vraiment les en blâmer? Certaines facultés centrent délibérément l'apprentissage sur les notes, comme si l'objectif de l'apprentissage était la réussite d'un examen interne ou externe à la faculté. Pourquoi l'examen ne serait-il pas plutôt un moyen de vérifier l'atteinte des objectifs? Si une trop grande part de l'activité en classe est centrée sur les

notes, le conflit de rôles où se trouve le professeur, s'amplifie. Le professeur qui croyait être un facilitateur de l'apprentissage est devenu juge. Il peut même être soumis à un certain chantage, quand il est lui-même évalué par ses étudiants à des fins administratives. Il n'est pas rare non plus de voir des étudiants qui évitent d'émettre des opinions différentes de celles du professeur ou, tout simplement, qui n'osent pas lui dire qu'ils n'ont guère compris ses explications, de peur que cela ne leur porte préjudice.

Un autre des sérieux problèmes engendrés par la notation est celui des lacunes dans la formation d'un étudiant. Expliquons-nous: un étudiant peut échouer un test portant sur telle partie du cours ou tel acquis, compenser ses faiblesses par un bon résultat à un autre test, obtenir la moyenne requise et être promu. Ce type de situation fait souvent dire au professeur que tel étudiant ne devrait pas être promu, mais que faire? Cela peut même conduire à des situations aberrantes où la note obtenue à un examen théorique compense pour un stage raté, même si, de fait, aucun cours théorique ne peut vraiment suppléer aux apprentissages visés par le stage.

Les codes utilisés pour la notation n'ont pas tous la même signification et sont donc d'interprétation difficile. Les notes peuvent être attribuées à partir des résultats obtenus dans un ou plusieurs examens dont les valeurs métrologiques (fiabilité, validité) n'ont pas toujours été vérifiées au préalable. Si certains professeurs accordent des notes selon l'assiduité au cours, d'autres cotent l'effort de l'étudiant, d'autres, enfin, mesurent l'atteinte des objectifs. Il n'y a pas uniformité dans l'attribution des notes. Pour tel professeur, il n'y a que Dieu qui ait 100%, le professeur qui puisse avoir 90%, le meilleur étudiant ne peut alors espérer que 80% comme note maximale. Pour un autre, l'échelle réelle s'étend bien jusqu'à 100. Les conventions qui régissent l'attribution des notes sont rarement exprimées de façon claire, encore qu'on ne puisse assurer que ces conventions soient respectées.

D'aucuns déplorent la subjectivité des professeurs. De là à voir dans l'examen objectif la solution à tous ces maux... Pourtant, dans ce cas, si la correction est objective, le choix et la formulation des questions n'en sont pas moins subjectifs. La plupart du temps, qu'il s'agisse de tests, objectifs ou non, de dossiers scolaires, émis ou non par ordinateur, la notation automatise la prise de décision; c'est-à-dire que la note de passage, par exemple, rend plus mécanique la prise de décision concernant les promotions.

De plus, l'évaluation normative, souvent utilisée pour attribuer des notes, présente l'inconvénient suivant: le succès ou l'échec d'un étudiant dépend alors des autres étudiants de sa classe. Selon que l'étudiant fait partie d'un groupe «fort» ou d'un groupe «faible», il a plus ou moins de chances de réussir.

Soulignons aussi que nos systèmes de notation contribuent à créer un climat d'anxiété et un esprit de compétition. En effet, les notes motivent les étudiants qui ont un esprit compétitif. De plus, diverses études indiquent que ces systèmes ne favorisent pas la démocratisation de l'enseignement: si

l'enseignement est accessible à tous, ce sont les mêmes classes sociales qu'auparavant qui y sont favorisées. Les étudiants les plus doués pour affronter nos systèmes de notation sont issus des classes sociales privilégiées.

Devant tous ces problèmes associés à la notation, certains professeurs les jugent insolubles et se désintéressent de tout ce qui s'y rapporte. Ils ont perdu confiance en la valeur de ce qu'ils font lorsqu'ils accordent des notes et ne veulent plus poser de jugements rigoureux sur les acquis des étudiants. Parfois, ils utilisent des moyens comme l'attribution d'une même note à tout le monde, sans plus de vérification. Ces pratiques favorisent le phénomène appelé «l'inflation des notes».

La plus lourde des conséquences demeure tout de même la dissociation trop fréquente entre l'apprentissage et l'évaluation. Les fonctions réservées à l'évaluation dans nos institutions sont plus souvent celles de sanction, de certification, plutôt que celle de support à l'apprentissage.

Nous constatons que plusieurs facultés ayant fait la sélection de leurs étudiants au moment de l'admission se servent tout de même d'une évaluation qui a toutes les caractéristiques de celle utilisée à des fins de sélection, alors que seulement quelques rares étudiants présentant des problèmes d'apprentissage majeurs ne sont pas promus d'une année à l'autre.

A quoi servent les notes?

Nous avons vu précédemment que les notes servaient à des fins administratives, c'est-à-dire à l'admission, à la promotion, à la sélection, à l'attribution de bourses, etc. Considérons quelques-uns de ces rôles et étudions l'éventualité d'abolir les notes.

D'abord les notes servent à l'admission à l'université. On croit donc que les succès scolaires au niveau collégial constituent une garantie valable du succès au niveau universitaire. Cependant, les critères de sélection sont parfois faibles. Par exemple, les cours de mathématiques, exigés pour être admis à des facultés comme Education physique, Médecine, etc., ne sont peut-être pas les meilleurs prédicteurs de succès dans ces domaines. L'abolition des notes au niveau collégial obligerait les universités à trouver d'autres moyens de sélection; on peut juger une telle mesure embarrassante. Cependant, les procédés utilisés pour sélectionner les étudiants à l'admission sont de plus en plus remis en question. Pourquoi ne verrait-on pas là l'occasion de chercher des solutions nouvelles?

Si nous considérons la promotion d'un cours à un autre, d'une année universitaire à une autre, nous constatons que plusieurs facultés ayant fait leur sélection à l'admission et ne faisant qu'éliminer les quelques rares étudiants présentant des problèmes d'apprentissage majeurs, ont tout de même conservé une évaluation centrée sur la sélection. Certaines facultés utilisent même l'artillerie lourde de la docimologie (soit dit en passant, nous n'avons rien contre celle-ci, nous sommes plutôt favorable à son bon usage) pour améliorer l'évaluation normative.

N'y aurait-il pas mieux à faire? L'évaluation ne devrait-elle pas plutôt favoriser une meilleure compétence chez un plus grand nombre d'étudiants?

On objectera peut-être que les notes permettent à l'employeur de choisir les meilleurs candidats qui sortent de l'université. Pourtant, bon nombre d'employeurs ne font plus confiance aux notes et se contentent plutôt du diplôme. D'autres, ne croyant pas que le succès académique entraîne le succès sur le marché du travail, ne se préoccupent pas du tout des notes. Pour eux, c'est la société qui évaluera les diplômés. C'est la réussite au cours de la période de probation ou d'entraînement qui sera le meilleur indice. Des recherches diverses confirment ces hypothèses. Du côté des carrières professionnelles, qui a jamais demandé à son médecin ou à son avocat s'il a obtenu son diplôme à 85% ou à 75%? Les notes ne seraient-elles valables qu'à l'université?

Un autre argument souvent invoqué en faveur des notes veut que celles-ci soient un facteur de motivation pour les étudiants. Quelle est la valeur de cet argument? Ce renforcement extrinsèque provoque-t-il chez les étudiants des attitudes que nous valorisons?

Voie nouvelle à explorer

Que faire, si nous songeons aux maigres avantages offerts par notre système de notation, pour compenser toutes les difficultés engendrées par celui-ci? Devons-nous tenter de l'améliorer par des règlements? Devons-nous changer notre système de notation pour un système de lettres? La formule succès/échec serait-elle la solution? A notre avis: non! On verrait surgir d'autres problèmes semblables. Ainsi, on a vu apparaître des A+, A++, A-, ... ces codes n'étant, très souvent, que la transcription d'un certain pourcentage (par exemple: A++ pourrait correspondre aux pourcentages entre 95 et 100). La formule succès/échec est aussi controversée. Plusieurs demeurent sceptiques quant à ses effets positifs sur la motivation à apprendre.

Évidemment, nous pouvons tenter d'atténuer les effets négatifs de notre système de notation et chercher à l'améliorer en vérifiant, par exemple, les qualités métrologiques des instruments utilisés.

Mais, ne serait-il pas temps de penser à autre chose? Ne pourrions-nous pas songer à abolir les notes? L'évaluation ne pourrait-elle pas jouer un rôle différent dans notre système éducatif?

Il ne faudrait pas croire que l'abolition des notes entraînerait la disparition de l'évaluation dans l'enseignement. Il pourrait quand même y avoir évaluation, mais celle-ci jouerait alors un tout autre rôle que celui qui lui est habituellement dévolu. Une évaluation qui, au lieu d'être centrée sur la sélection, favoriserait un meilleur apprentissage ne serait-elle pas plus appropriée? Le professeur pourrait être plus exigeant en ce qui concerne la formation des étudiants, en leur demandant, par exemple, d'atteindre certains niveaux de compétence. Mais il pourrait compenser cette plus grande exigence en permettant à l'étudiant de se reprendre lorsqu'il commet certaines erreurs dans son apprentissage. C'est un feedback qui s'accompagne d'une revérification des acquis. Nous parlons alors d'évaluation critériée et formative.

L'évaluation devient ainsi plus démocratique en ce sens qu'elle permet à plus d'étudiants d'atteindre les niveaux de compétence requis. Les étudiants qui éprouvent des difficultés dans le système actuel ont certaines chances de s'en tirer: on ne laisse plus les difficultés se perpétuer, s'amplifier jusqu'à la catastrophe, c'est-à-dire l'échec sans possibilité de corriger ses faiblesses ou la promotion sans les acquis nécessaires. La solution résiderait peut-être dans une modification profonde des règles qui régissent notre système d'éducation. Divers moyens peuvent nous y aider.

La précision des objectifs d'apprentissage facilite l'évaluation, puisque ces objectifs peuvent servir de critères. De plus, la rétroaction est plus significative si elle est reliée à chacun des objectifs. Lorsque les étudiants savent ce qu'on attend d'eux, ils évitent bon nombre d'activités non pertinentes, comme celle de deviner ce qui est important pour le professeur.

Une évaluation centrée sur l'apprentissage plutôt que sur les notes tend à diminuer sérieusement la compétition entre les étudiants. Ceux-ci, ne l'oublions pas, veulent apprendre et non uniquement obtenir des notes. Les étudiants étant plus concernés par leur apprentissage y trouvent une motivation intrinsèque.

L'évaluation formative repose sur un changement profond des attitudes de l'enseignant et de l'enseigné. Conçue comme une aide à l'apprentissage, elle permet à l'étudiant d'être plus responsable de son travail. Cependant, il n'est pas facile de créer un climat qui incite, par exemple, un étudiant à demander des explications à un camarade ou au professeur sur une question à laquelle il a répondu avec succès, mais par hasard ou avec une connaissance partielle du sujet. Est-il en mesure de se rendre compte qu'il est le seul pénalisé lorsqu'il cache ses difficultés? Le professeur est le moteur de ce changement d'attitudes. Lui-même, se voit attribuer un rôle différent: il aide l'étudiant à parfaire son apprentissage, à atteindre les objectifs.

Une évaluation ainsi conçue permet d'exiger de l'étudiant qu'il atteigne une compétence minimale spécifique ou une maîtrise parfaite des apprentissages pour pouvoir inférer qu'il a réussi son cours. Le succès dans un secteur ne compense pas pour un échec dans tel autre si les deux secteurs sont jugés essentiels à la formation. Il est possible alors de s'assurer qu'il n'y a pas de lacunes dans la formation. Certains objectifs habituellement négligés peuvent être l'objet d'une évaluation: par exemple, il est plus sûr que l'on puisse rendre les étudiants capables de s'auto-évaluer, puisque cet apprentissage n'est pas biaisé par une note à accorder.

Conclusion

Ce mode d'évaluation est-il farfelu? innovateur? plausible? À vous d'en juger. Certaines expériences ont été tentées dans ce domaine et se sont révélées positives. On y a constaté l'amélioration des acquis des étudiants et le développement d'attitudes souhaitées mais négligées par les programmes d'études. Cependant, une telle évaluation se vit habituellement à l'intérieur des contraintes du système de notation: il faut donc chercher des moyens pour concilier ce mode d'évaluation avec les exigences administratives.

Ne serait-il pas temps de s'interroger sur les possibilités d'un système sans notes? Evidemment, divers points restent à préciser dans cette voie. Il y existe aussi des difficultés, mais celles-ci ne seraient-elles pas mineures comparativement à celles que pose notre système de notation, difficultés étudiées depuis fort longtemps mais auxquelles on a proposé peu de solutions satisfaisantes jusqu'à maintenant.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 1 No 3, Sept. 1976, pp. 15-17.

Section 5

Des expériences en évaluation

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

Un système informatique de contrôle de l'apprentissage

Gilles Génier

Depuis septembre 1971, nous appliquons, à l'Ecole Polytechnique, un système informatique de contrôle de l'apprentissage (S.I.C.A.). Plus de 500 étudiants suivent un cours de statique de l'ingénieur contrôlé par ordinateur! La statique est aussi fondamentale pour un étudiant de Polytechnique que l'anatomie l'est en médecine ou la comptabilité aux H.E.C.

La sonate ou les gammes?

Depuis longtemps, l'enseignement de la statique causait bien des difficultés. Entre 1965 et 1971, les professeurs concernés y ont apporté plusieurs améliorations, mais le problème persistait. Ils ont alors identifié le problème à peu près comme suit:

- La préparation des étudiants qui arrivent à Poly diffère d'un étudiant à l'autre. Les collèges et les cégeps poursuivent des objectifs de formation générale similaires, mais la composante scientifique de cette formation varie très largement d'une institution à l'autre, tant en physique qu'en mathématiques.
- Les bases de la statique de l'ingénieur sont très peu attirantes pour l'étudiant qui suit, au même moment, des cours d'initiation à des sciences plus prestigieuses. On doit cependant lui demander d'acquérir une certaine habileté dans les calculs de base avant de le laisser s'attaquer à des calculs relatifs à des structures de machine, d'édifice ou de pont, ou à tout autre arrangement mécanique. De la même façon, on exige de l'étudiant pianiste une habileté dans les gammes, même s'il préfère s'attaquer à une sonate de son choix.
- Les professeurs qui reçoivent nos étudiants se plaignent de leur incompétence en ce qui a trait aux calculs simples (les gammes), tout en appréciant leur connaissance de plusieurs méthodes complexes de calcul (les sonates). L'étudiant connaît la sonate; il peut la pianoter, la chantonner ou en parler agréablement, mais il est incapable de l'exécuter au piano convenablement.

- Pour une bonne majorité de nos étudiants, la statique n'est pas une matière de formation mais une matière fondamentale, suffisamment importante pour qu'on doive exiger d'eux une certaine habileté dans la détermination des forces. Cette habileté peut ensuite être transposée dans d'autres domaines vectoriels sans relation avec la statique.
- Le groupe de professeurs a déjà essayé, depuis plusieurs trimestres, toutes sortes de moyens tendant à améliorer le rendement des étudiants: présentations au moyen d'acétates, devoirs et séances d'exercices dont la notation affectait ou non la note globale de l'étudiant. Chaque effort nouveau des professeurs est bénéfique, mais son effet est toujours insuffisant et passager.
- Des questionnaires remplis par les étudiants montrent assez clairement que la qualité des professeurs n'est pas en cause. Les cours sont bien préparés bien présentés et les examens sont acceptables. La matière paraît facile aux étudiants, mais, malgré les avertissements répétés des professeurs, ils ne travaillent pas suffisamment et éprouvent beaucoup de difficultés aux examens. De 35 à 40 pour cent des inscrits échouent malgré la générosité des professeurs lors de la correction des copies d'examen.

... les gammes!

En 1971, l'Ecole Polytechnique décide d'organiser, sous les auspices de l'American Society for Engineering Education, le colloque annuel sur les méthodes d'enseignement pour la section française de la région du Saint-Laurent. On y a parlé abondamment d'objectifs de comportement et d'individualisation de l'enseignement. Je me suis alors laissé séduire par les possibilités qu'offraient les méthodes d'individualisation autant que par l'enseignement par objectifs, peu à la mode dans les universités du Québec, à l'époque. Je réussis à communiquer mon enthousiasme au groupe de professeurs de statique et nous décidons de nous mettre à l'oeuvre.

Après avoir lu des comptes rendus d'une dizaine d'expériences d'individualisation, nous arrivons à la conclusion que l'individualisation intégrale, enseignement et contrôle, est difficilement applicable à des groupes nombreux. Au surplus, les «grands groupes» dont on parle dans ces publications comptent de 20 à 25 étudiants. Nous en attendons entre 400 et 500. La plus grande difficulté réside dans le contrôle individuel de l'étudiant qui progresse à son propre rythme. Il nous faut donc nous attaquer à ce problème tout particulièrement.

Même pour un groupe restreint, l'individualisation de l'enseignement entraîne un problème de contrôle. Lorsque l'étudiant progresse en effectuant son travail personnel, il faut multiplier les contrôles pour qu'il soit à même d'évaluer le rendement de ses efforts, et ce, le plus fréquemment possible. A plus forte raison, pour ce qui est d'un très grand groupe, la préparation et la correction des contrôles deviennent une corvée insupportable pour le professeur. En dépit des problèmes qui s'annoncent, nous élaborons une solution (SICA) et nous fonçons. Aujourd'hui, je trouve cette attitude naïvement trop

agressive. Nous aurions dû attendre que SICA soit fonctionnel et que notre matériel didactique soit prêt, plutôt que de foncer dans la classe avec un nouveau cours partiellement préparé.

Je passe outre aux détails pénibles de nos débuts alors que ni le Service pédagogique de l'Ecole Polytechnique, ni celui de l'Université de Montréal n'existaient encore. En gros, après beaucoup d'efforts et de travail, sans support technique raisonnable, nous avons mis sur pied, en six douloureux mois, un cours complètement individualisé (enseignement et contrôles) et qui élimine même les limites du trimestre. Nous y recevons plus de 500 étudiants par trimestre. Ce cours s'appuie sur un système de contrôles des connaissances de l'étudiant, SICA, dont nous sommes très fiers aujourd'hui. Depuis ce temps, SICA a été grandement amélioré grâce à des fonds de la compagnie IBM Canada Ltée. Les paragraphes qui suivent décrivent quelques-unes des caractéristiques de notre méthode.

Rythme de l'étudiant

Le rythme de l'étudiant est respecté: le temps qui lui est alloué pour traiter un sujet donné, tout comme le moment qu'il aura choisi pour y travailler. Il n'y a aucun horaire obligatoire. Même si, dans le plan du cours, des modules sont prérequis à d'autres, ce plan prévoit plusieurs cheminements plus ou moins indépendants, ce qui augmente encore la liberté de choix de l'étudiant.

Il en est de même pour les contrôles. L'étudiant peut non seulement venir subir des contrôles à la date de son choix, mais il est, de plus, libre de passer ses trois heures sur un seul module, tandis que d'autres essaieront deux, trois et même jusqu'à cinq modules durant la même période de contrôle. Nous avons mis au point un système SICA-I (I pour «Interactif») où même l'horaire des contrôles pourrait disparaître. Ce système n'est pas implanté pour des raisons budgétaires, mais, avec la réduction des coûts des terminaux, il est probable que ce système deviendra économiquement réalisable un jour.

Après deux ans d'essai, nous avons fait une entorse à la liberté des étudiants, parce qu'ils en abusaient. Ils sont maintenant soumis à un système de pénalités qui accorde, à celui qui fait son travail en deux ou trois trimestres, une moins bonne note qu'à celui qui le fait en un seul trimestre.

Maîtrise du sujet

Pour favoriser le rythme personnel de l'étudiant, il faut structurer la matière, encore plus que s'il s'agissait d'un cours magistral. Le cours est divisé en 19 modules. Certains sont prérequis à d'autres, mais le plan de cours n'est pas linéaire: on y trouve 4 embranchements. L'étudiant est soumis à une évaluation continue et on exige de lui la maîtrise du sujet. Cette maîtrise du sujet correspond à des normes minimales de passage variant entre 75 et 85 pour cent. L'évaluation est continue puisque, tant qu'un étudiant ne maîtrise pas un module donné, il ne peut s'attaquer au suivant. Cette condition ne s'applique pas pendant une même période de contrôle, puisque l'étudiant ne sait pas instantanément s'il a réussi un module, et qu'il peut essayer d'en réussir un autre.

L'étudiant a droit à autant de reprises qu'il le désire sans autre pénalité que le retard qu'il prend dans son travail. Seules ses réussites comptent pour sa note, ses échecs lui indiquant ses points faibles. Comme tous ceux qui réussissent un module maîtrisent en principe le sujet, ils reçoivent tous la même note pour chaque module. On distingue les étudiants entre eux par la quantité de modules maîtrisés. La réussite du bloc de modules qui constitue la base essentielle à connaître pour tout ingénieur, donne la note minimale de passage. L'étudiant choisit lui-même d'améliorer plus ou moins cette note en maîtrisant un certain nombre de modules optionnels et il peut poursuivre son travail au trimestre suivant, s'il le désire.

L'étudiant est guidé dans son travail par:

- un cahier de travail qui comprend un plan de cours, des objectifs de comportement détaillés, des exemples, des remarques, des directives, des problèmes à résoudre, des exemples de questionnaires d'examen, etc.;
- des directives générales enregistrées sur cassettes qui précèdent une série de solutions complètes à des problèmes typiques que l'étudiant doit tenter de résoudre avant d'en vérifier la solution;
- le professeur qui, disponible tous les après-midi, joue le rôle d'un expert-conseil que l'étudiant vient consulter au besoin.

Cette formule offre de très grandes possibilités sur le plan des moyens audio-visuels que nous n'avons pas encore exploitées.

L'ordinateur est de corvée

Pour réussir un tel projet, il faut assurément une utilisation optimale de l'ordinateur, sans toutefois créer trop de dépendances envers le centre de calcul. Il faut donc lui refiler toutes les corvées possibles. Dans notre projet, c'est le centre de calcul qui assume les fonctions suivantes:

- conservation de notre banque de questions;
- préparation de la majorité des questionnaires de contrôle (environ 8000 questionnaires différents par trimestre);
- correction de ces questionnaires et de ceux qui sont conçus par les professeurs;
- mise à jour des dossiers de chaque étudiant en vérifiant, dans chaque cas, les prérequis;
- calcul de la note de chaque étudiant compte tenu des pénalités dont nous avons parlé précédemment;
- surveillance des cas de plagiat entre deux étudiants;
- préparation de données statistiques pour la réussite de chaque question de la banque, et de chaque module du cours.

Bien sûr, une secrétaire surveille de très près ce travail en examinant tous les rapports qu'il nous imprime. Elle doit consacrer environ quinze heures par semaine à cette tâche.

La préparation de questionnaires, tous différents, pourrait nous amener des plaintes quant aux différences de difficulté d'un questionnaire à l'autre. Chaque questionnaire est conçu suivant des normes, de telle sorte qu'ils sont tous globalement de même difficulté et qu'ils couvrent tous chaque objectif du module. Ainsi nous n'avons jamais eu de plainte à ce sujet.

Satisfaction chez les étudiants

Pendant les deux premiers trimestres, nous avons considéré avec soin toutes les demandes des étudiants et nous avons toujours répondu, favorablement ou non, avec rapidité à ces demandes. Les étudiants l'ont apprécié et, lorsque nous leur avons distribué des questionnaires d'évaluation beaucoup plus complexes que de coutume, ils ont très bien collaboré avec le «comité de surveillance» qui évaluait l'expérience. Les résultats étaient généralement favorables, ce qui contrastait avec le passé. Les étudiants les plus forts étaient très enthousiastes, l'étudiant moyen était satisfait et une faible minorité aurait préféré des cours traditionnels. Depuis, nous n'avons pas tenté de réévaluer la satisfaction de l'étudiant à l'aide de questionnaires, mais, à quelques reprises, des rapports d'étudiants, à la suite des semaines d'éducation, ont cité à plusieurs reprises le cours de statique par exemple. Par contre, d'autres interventions révèlent un noyau restreint d'insatisfaits. La présentation de ce cours n'est pas sans défaut, mais pour les corriger, il faudrait y mettre une certaine quantité d'énergie que les professeurs ne semblent pas prêts à déployer.

... et chez les professeurs

Il est certain que les professeurs ne sont pas tous d'accord sur les avantages et les inconvénients de cette méthode d'enseignement. Voici ce que j'entends le plus fréquemment quant aux avantages:

- Ce cours demande au professeur de très bien connaître la matière dans son ensemble mais n'exige aucune préparation immédiate.
- Un nouveau professeur s'adapte facilement à cette méthode d'enseignement.
- Cette méthode ne demande aucune qualité spéciale comme, par exemple, pouvoir susciter l'intérêt de 60 étudiants plus ou moins réceptifs au moment où la cloche sonne.
- Les contacts avec les étudiants sont nombreux et personnalisés.
- Cette méthode permet d'utiliser tous les media disponibles (malheureusement ce n'est pas encore fait dans notre cours.)

Par contre, j'entends aussi parler d'inconvénients:

- Le professeur répète fréquemment la même explication, ce qui devient monotone.
- Il résout des petits problèmes particuliers, mais ne peut présenter de synthèses qui mettent en valeur sa compétence.
- Il ne se sent pas très impliqué dans la préparation du matériel didactique, bien que sa participation active soit très appréciable et souhaitée.

- Il peut utiliser le système sans rien y apporter. Du point de vue du système, certains professeurs ne sont que des «clients».
- Il a plus de contacts avec les étudiants médiocres ou mésadaptés qu'avec les meilleurs.

Et si c'était à refaire?

Si c'était à refaire, je n'hésiterais pas, mais j'accorderais une plus grande importance, au départ, à des conditions qui nous paraissent maintenant essentielles:

- une équipe adaptable et décidée;
- une aide technique assurée en pédagogie et en audio-visuel;
- un budget «officiel»;
- une organisation matérielle sans faille.

Quant à la méthode à utiliser pour individualiser l'enseignement, j'y ajouterais (et c'est très facile à faire dans notre système) un choix d'activités pour l'étudiant. Les sujets les plus difficiles seraient présentés de plusieurs façons.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 2 No 2, Avril 1977, pp. 7-10.

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

La banque d'objectifs et d'items : un exemple d'application

Huguette Bernard

Serge Normand

L'évaluation des apprentissages scolaires demeure un processus complexe qui s'inscrit souvent comme un mal nécessaire autant pour le professeur que pour l'étudiant. L'évaluation reste pourtant partie intégrante de l'acte éducatif, dans la mesure où les agents impliqués identifient eux-mêmes des objectifs précis d'apprentissage et reconnaissent la nécessité de vérifier l'atteinte de ces objectifs. Une des stratégies pour améliorer la qualité de l'évaluation des apprentissages et en diversifier les résultats peut être la création d'une banque d'items jumelés à une banque d'objectifs.

La définition d'une banque d'items[1] réfère à une collection de questions rassemblées selon un ordre quelconque et caractérisée par un système qui permet de manipuler efficacement et rapidement chacun des éléments d'information qui la constituent.

Il est fréquent de compter plusieurs centaines d'items dans une banque. Les items sont destinés à composer des questionnaires d'évaluation mesurant des objectifs et des contenus spécifiques.

Une situation concrète

Il y a plus d'un an, nous entreprenions avec le département de Gynécologie-obstétrique de la Faculté de Médecine de l'Université de Montréal un projet de banque d'objectifs et d'items dont le but premier était de répondre à un besoin d'évaluation formative. Ce projet, identifié sous le nom de *gynobst*, favorisa une meilleure définition des objectifs d'apprentissage et une exploitation rationnelle d'un bassin de questions existantes.

L'évaluation formative a pour but de renseigner l'étudiant sur ses performances par rapport aux objectifs à atteindre. Les résultats, qui ne sont pas consignés au dossier des étudiants, doivent leur être communiqués rapidement pour favoriser un retour aux objectifs non maîtrisés. De plus, la forme et le contenu des tests doivent les préparer à leur examen de certification venant d'un organisme externe.

Le département doit, par ailleurs, disposer d'un bassin d'au moins 1600 questions pour répondre chaque année à l'une des formules d'évaluation pratiquée dans son programme de stages. Cette formule consiste à administrer, au cours d'une période de douze semaines, 4 examens distincts d'environ 100 questions chacun. Quatre stages semblables sont répétés annuellement et touchent une population d'environ 200 étudiants.

La mise sur pied de la banque d'objectifs et de la banque d'items s'est réalisée en quatre étapes:

- Elaboration d'un système de classification commun aux objectifs et aux items.
- Définition et reformulation des objectifs d'apprentissage.
- Rédaction et révision de plus de 1000 items.
- Elaboration d'un logiciel informatique approprié.

Le système de classification

L'efficacité d'une banque d'items repose sur un système de classification rigoureux qui caractérise et les objectifs et les items.

Onze (11) paramètres de classification ont été retenus dans le cas du projet *gynobst*. Certains de ces paramètres sont communs à la fois aux objectifs et aux items et permettent de détailler de façon très sophistiquée les contenus mesurés. D'autres appartiennent uniquement aux items et apportent une spécificité complémentaire.

L'exemple suivant illustre l'énoncé d'un objectif et les paramètres de classification qui s'y rattache.

Sans pour autant décrire tous les détails de la classification, mentionnons que l'objectif cité plus haut se rapporte à la section *obstétrique pathologique* (Se 3) couvrant le module *pathologie du 3e trimestre de la grossesse* (Mo 2) et porte sur le contenu spécifique *le saignement du 3e trimestre* (Co 4). Les quatres derniers paramètres caractérisent l'objectif.

Les paramètres propres à l'item identifient, entre autres, la clientèle à qui s'adresse la question et en fournissent la bonne réponse.

Énoncé d'un objectif :	Faire le diagnostic différentiel d'un saignement du 3e trimestre
Paramètres de classification de l'objectif :	Se - Mo - Co - Ot - Dt - Nt - No 3 2 4 1 1 53 1

Paramètres de l'item :	Se - Mo - Co - Ot - Dt - Nt - No - Et - Mc - Vr - Br 3 2 4 1 1 53 1 1 11 1 3
Énoncé d'un item :	Lorsqu'une patiente, au troisième trimestre, présente une hémorragie utérine indolore, on doit d'abord penser à 1. un décollement prématuré du placenta normalement inséré; 2. une dyscrasie sanguine; 3. un placenta praevia, 4. une lésion sanguinolente du col; 5. une hémangiome vaginal.

Les objectifs d'apprentissage

Un système de classification aussi détaillé a favorisé une nouvelle formulation des objectifs d'apprentissage. Pour ce faire, nous nous sommes interrogés sur les activités à maîtriser dans le stage et, de là, nous avons créé notre propre taxonomie. Cette classification des objectifs correspond au cheminement d'un étudiant auprès d'un patient.

	0	Prérequis
Connaissance	1.0	Connaître
Compréhension	2.0	Comprendre
Application	3.3	Faire le questionnaire approprié
	3.6	Faire l'examen approprié
Analyse	4.2	Rédiger l'histoire de cas et l'examen physique
	4.4	Faire le diagnostic différentiel
Synthèse	5.2	Faire l'investigation
	5.3	Interpréter les résultats
	5.5	Faire le diagnostic final
	5.7	Faire le traitement
Évaluation	6.0	Faire le pronostic

Vingt-cinq médecins du département de Gynécologie-obstétrique ont été mis à contribution pour ce projet. Ils ont été libérés à raison de trois heures par semaine pendant un an. Les trois premiers mois ont été consacrés à la redéfinition des objectifs; les mois suivants, à la création, correction et codification des questions.

La validation des items

La validation des items s'est faite à partir d'un bassin de plusieurs centaines de questions déjà existantes. Au fur et à mesure que l'on révisait les questions, on les codifiait en fonction des paramètres de classification retenus.

De nouveaux items ont dû être rédigés pour répondre aux besoins des objectifs et du contenu. Notre objectif final était d'atteindre 1000 questions au cours de la première année. Après un an, la banque d'items contient 1078 questions.

Après la rédaction, venait l'expérimentation. Aussitôt qu'une équipe avait terminé son quota de questions, celles-ci étaient expérimentées auprès de groupes d'étudiants de médecine.

De nouveaux objectifs et items pourront à tout moment être insérés dans la banque pour mieux répondre au besoin du curriculum et de l'évaluation formative. Deux types de questions à choix multiple ont été exploités pour cette banque d'items: le type A, composé d'un énoncé affirmatif ou interrogatif et d'un choix de 5 réponses dont 4 leurres,[2] et le type K permettant à l'étudiant de choisir parmi des combinaisons variées de réponses.

Les exemples suivants sont tirés de la section gynécologie générale.

Type A Enoncé: Quelle est la maladie qui forme un ulcère vulvaire serpigineux, superficiel et granuleux?

Choix de réponses:

1. Le condylome accuminé.
2. Le granulome inguinal.
3. Le lymphogranulome vénérien.
4. La syphilis.
5. La leukoplasie.

Type K Enoncé: L'endométrie aiguë se manifeste par

a) de la fièvre,
b) un utérus engorgé, sensible,
c) un écoulement cervical purulent,
d) une douleur à la fosse iliaque gauche.

Choix de réponses:

1. a et b et c sont vrais.
2. a et c sont vrais.
3. b et d sont vrais.
4. d est vrai.
5. tous les énoncés sont vrais.

Le logiciel informatique

L'ordinateur s'est révélé un outil indispensable pour manipuler rapidement les quantités imposantes de données reliées au projet *gynobst*. Un logiciel informatique a été développé pour faciliter les opérations suivantes:

- le stockage et la mise à jour du fichier d'objectifs et du fichier d'items;
- la sélection des items et la préparation des questionnaires d'évaluation;
- la correction des examens et la production des analyses statistiques pertinentes;
- la présentation d'un rapport d'évaluation aux étudiants, aux professeurs et aux responsables du département.

Les ressources offertes par les ordinateurs *cyber* du Centre de calcul de l'Université de Montréal permettent d'effectuer les opérations de stockage et de préparation des questionnaires en mode conversationnel. Ce qui signifie qu'aucun apprentissage d'un langage informatisé n'est requis de la part des usagers pour ces opérations.

A l'aide d'un terminal,[3] les responsables du projet *gynobst* peuvent très facilement ajouter, modifier et éliminer des objectifs ou des items de la banque.

De même, le professeur peut commander un questionnaire d'évaluation selon des schémas prédéfinis ou encore le composer en quelques minutes en variant la combinaison des paramètres de classification. D'une part, des tableaux de spécification ont été définis pour les quatre examens formatifs administrés au cours d'un stage: un examen portant sur les prérequis, deux examens partiels et un examen global. Ces quatre examens suivent l'évolution de l'apprentissage de l'étudiant. D'autre part, il demeure toujours possible de commander de courts tests mesurant, par exemple, un module particulier ou couvrant un niveau taxonomique particulier.

Le texte des questions, imprimé directement au terminal, peut ensuite être reproduit en plusieurs exemplaires et distribué aux étudiants. Ceux-ci répondent sur des formulaires traitables par lecteur magnétique, ce qui assure un retour rapide à l'ordinateur pour les opérations correction et rapport d'évaluation.

Le système exploité présentement pourrait permettre de produire un questionnaire différent qui répondrait au besoin de chacun. S'il était nécessaire, le système pourrait facilement être modifié pour permettre à l'étudiant d'individualiser sa requête d'évaluation, et ce, au terminal même.

Le rapport d'évaluation

Tel qu'exploité dans le cadre du projet *gynobst*, la banque d'objectifs et d'items assure à chacune des instances impliquées (étudiant, professeur, département) un feedback particulier.

Chaque *étudiant* reçoit des indications personnelles, en termes d'objectifs maîtrisés ou à maîtriser, dans un laps de temps très court, pour lui permettre un retour sur les contenus qui soulèvent des problèmes.

Le *professeur* peut facilement identifier les objectifs qui ne sont pas atteints et décider de privilégier des contenus particuliers. Il lui est évidemment très facile de commander, à n'importe quel moment, un questionnaire pour mesurer un ou des contenus très particuliers.

Enfin, les *responsables* du département obtiennent, à partir des analyses statistiques effectuées, des renseignements sur la qualité des apprentissages des étudiants et peuvent en retirer, en temps utile, des indications qui permettent de modifier, au besoin, le programme d'étude.

Mentionnons à la toute fin que les paramètres statistiques retournent aux équipes de travail qui peuvent juger de la qualité des items.

Le feedback

En résumé, la création d'une banque d'objectifs et d'items aura permis au département de Gynécologie-obstétrique de récupérer un matériel d'évaluation déjà existant; de favoriser un retour aux objectifs de son curriculum et de mettre à contribution la collaboration de vingt-cinq médecins de quatre centres hospitaliers universitaires pour un projet commun.

Les premières expériences d'exploitation de la banque démontrent que ce système intégré aux autres éléments du curriculum permet à l'étudiant de mieux saisir les objectifs de son programme d'étude tout en l'encourageant à prendre en main son auto-évaluation. L'intérêt manifesté par les professeurs et par les étudiants dans l'utilisation du feedback porte à croire que ce mode d'évaluation à caractère formatif peut contribuer à faire de l'évaluation un instrument réel d'apprentissage.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 3 No 4, Déc. 1978, pp. 21-24.

1. En évaluation «item» est synonyme de «question».

2. L'anglicisme «distracteurs» est à proscrire.

3. Le terminal du type machine à écrire peut être installé dans un centre hospitalier et relié par ligne téléphonique au Centre de calcul.

SERVICE PÉDAGOGIQUE UNIVERSITÉ DE MONTRÉAL

L'évaluation de la formation clinique en ergothérapie

Francine Forland

Raymonde Hachey

En ergothérapie, tout comme dans la plupart des professions de la santé, la formation clinique des étudiants soulève plusieurs difficultés. Cet article veut présenter une ébauche de solutions concernant l'une de ces difficultés, soit l'évaluation de la formation clinique.

Pour fins d'information, soulignons que l'enseignement clinique en ergothérapie est assumé par les cliniciens (moniteurs de stage). Dans chaque centre, un professeur-conseiller est assigné comme personne-ressource et une évaluation écrite de l'étudiant est demandée au clinicien. Des rencontres bi-annuelles réunissent les coordonnateurs de formation clinique, les professeurs-conseillers et des représentants des étudiants.

Au cours des trois années de formation en ergothérapie, les périodes de stage sont réparties de la façon suivante: à la fin de la première année, un stage d'environ une semaine est planifié; en deuxième année, douze jours de stage sont prévus à chaque trimestre; en troisième année, on compte vingt-quatre jours de stage par trimestre. Ces différents stages font partie du curriculum de formation et des crédits leur sont alloués.

Etant donné l'envergure moins importante du stage de la 1re année, le travail présenté ici concernera seulement l'évaluation de la formation clinique des étudiants des 2e et 3e années.

La définition du problème

Jusqu'en janvier 1977, pour évaluer l'étudiant, on demandait au moniteur de stage de commenter différents aspects du fonctionnement du stagiaire (évaluation du patient, application des modalités thérapeutiques, relations interpersonnelles, etc.). A partir de ce tableau général, on cotait l'étudiant (échec, bien, très bien, excellent) et cette cote se transformait en note (50%, 65%, 75%, 85%).

Cette façon d'évaluer et d'être évalué soulevait de nombreuses critiques de la part des cliniciens et des étudiants. Tels que rapportés au cours d'une réunion de coordonnateurs de la formation clinique, tenue en mai 1976, les commentaires suivants des cliniciens et des étudiants résument bien la situation d'alors: «Le formulaire d'évaluation en usage laisse une très large part à la subjectivité du superviseur; aucun objectif spécifique à atteindre, aucune façon objective de coter l'étudiant, peu ou pas de participation de l'étudiant à son évaluation, aucune uniformité d'un centre à un autre quant à l'utilisation du formulaire».

L'insatisfaction était également vécue par les professeurs, puisque l'évaluation ne permettait pas de cerner les apprentissages spécifiques de l'étudiant et la note obtenue n'était pas nécessairement proportionnelle à cet apprentissage. Des objectifs spécifiques en fonction des principaux champs de pratique en ergothérapie avaient déjà été fournis aux cliniciens mais il était évident que ces objectifs n'étaient pas intégrés à l'enseignement donné à l'étudiant ni à l'évaluation.

Le problème était donc le suivant: comment évaluer en milieu clinique de façon plus objective, l'apprentissage de l'étudiant (évaluation formative) et comment déterminer la note correspondant à cet apprentissage (évaluation normative)?

La formation d'un comité

A la suite de cette réunion, un comité fut formé et mandaté pour élaborer un nouveau formulaire d'évaluation de la formation clinique pour les trois niveaux d'enseignement.

Composé de quatre cliniciens représentant les principaux champs de pratique en ergothérapie, de deux étudiants et de deux professeurs, ce comité a tenu depuis sa formation des réunions hebdomadaires. Des consultations fréquentes auprès du Service pédagogique ont été assurées par Mme France Fontaine, docimologue.

L'élaboration du formulaire

La première démarche du comité fut l'analyse bibliographique sur la formation clinique dans différents domaines de la santé et sur l'évaluation de cette formation.

Après cette cueillette de renseignements, les trois parties en cause (cliniciens, étudiants et professeurs) ont fait part de leurs besoins et de leurs attentes quant au nouveau formulaire. A partir de cet échange, on a défini, les trois buts du nouveau formulaire. Par ordre de priorité, il veut:(1) assurer un feedback régulier à l'étudiant tout au long de son apprentissage, favorisant ainsi un cheminement progressif vers une compétence professionnelle; (2) améliorer et uniformiser la formation clinique dans les différents centres; (3) permettre l'obtention d'une note nécessaire pour le dossier de l'étudiant.

Le comité a alors étudié différentes évaluations tirées de l'analyse bibliographique. L'évaluation préparée par l'association américaine des ergothérapeutes, soit le «Field Work Performance Report»[1] (FWPR) a particulièrement retenu l'attention du comité à cause de ses caractéristiques particulières.

Avec l'autorisation de l'association, le comité a alors utilisé le «FWPR» comme outil de travail.

L'élaboration du nouveau formulaire

Dans un premier temps, il y eut comparaison entre les objectifs remis antérieurement aux moniteurs de stage et ceux du «FWPR»; une adaptation de cette évaluation a par la suite été entreprise. On a réduit le nombre d'objectifs et changé le mode de cotation et de notation. Dans le «FWPR», chaque cote (rarement, occasionnellement, fréquemment, constamment) était pondérée et le total des points donnait la note finale. Compte tenu du premier but du nouveau formulaire (feedback fréquent à l'étudiant), le comité décida de coter chaque objectif comme suit: «atteint», «non atteint», «non applicable». Un chiffre précisait la semaine du stage pendant laquelle l'objectif était évalué. Un même objectif pouvait être évalué plus d'une fois pendant le stage.

Pour établir la note, les membres du comité, après consultation auprès des professeurs et des cliniciens, ont divisé les objectifs en deux catégories: objectifs obligatoires (jugés essentiels pour devenir un ergothérapeute compétent) et objectifs complémentaires (jugés non essentiels mais importants pour un futur ergothérapeute). Tous les objectifs obligatoires devaient être atteints pour réussir le stage: si un seul ne l'était pas, c'était l'échec (voir tableau 1).

Tableau 1 - Notation du formulaire

OBJECTIFS				
%	Atteints		Non atteints	
	●	○	●	○
50			1	
70	tous	aucun		
80	tous	10		
90	tous	15		

Légende

● = objectif obligatoire

○ = objectif complémentaire

Exemple: pour obtenir la note de 70%, l'étudiant doit atteindre tous les objectifs obligatoires mais n'est pas tenu d'atteindre les objectifs complémentaires.

Par la suite, le comité décida d'élaborer un même formulaire pouvant être appliqué dans les différents champs cliniques d'ergothérapie. Cela avait pour but d'éviter une spécialisation prématurée de l'étudiant et assurait une approche globale du patient ce qui correspond à la philosophie même de l'ergothérapie. En ce qui a trait aux niveaux d'enseignement, ce même formulaire pouvait s'appliquer aux étudiants de 2e et de 3e année et favorisait ainsi une gradation des objectifs d'une année à l'autre. Un formulaire différent a été conçu pour les étudiants de 1re année.

L'expérimentation du formulaire et les résultats obtenus

Durant le trimestre d'hiver 1977, une première expérimentation du nouveau formulaire a été mise sur pied. Un questionnaire, rempli par 72 étudiants de 2e et de 3e année et par 64 moniteurs de stage, nous ont fourni les renseignements suivants. 64% des moniteurs trouvaient la terminologie utilisée claire et la phraséologie compréhensible. L'exigence d'atteindre tous les objectifs obligatoires semblait peu réaliste à 48% des moniteurs à cause de la somme de travail requis. Paradoxalement les notes obtenues par les étudiants étaient très élevées, ce qui démontrait l'atteinte de nombreux objectifs complémentaires (voir Tableau 2).

Tableau 2

Comparaison des notes obtenues par les étudiants de 2e et 3e année

Note obtenue	2e		3e	
%	n	%	n	%
70	16	40	8	25
80	12	30	18	56.25
90	12	30	6	18.75
TOTAL	40	100	32	100

Légende

n = n. d'étudiants

L'ensemble des cliniciens mettait en doute la valeur de ces résultats critiquant l'absence de critères quant à la qualité de l'atteinte des objectifs et à la gradation de cette atteinte. Les objectifs cotés non applicables ressortaient comme étant de deux types: non applicables en fonction du milieu

(exemple: objectifs concernant la postcure dans un centre ne recevant que des patients en phase aiguë); non applicables en fonction du contexte (exemple: événements imprévus lors du stage).

Afin d'évaluer l'aspect formatif et normatif du formulaire, les réponses sur le feedback et la standardisation de la note ont été prélevées à partir du questionnaire.

Tableau 3

	Feedback continu pour fins d'apprentissage		Standardisation de la note	
	code *	%	code	%
Avant l'utilisation du formulaire	1	37	1	33.3
Après l'utilisation du formulaire	1	44.4	5	48.1

Les notes 1 et 5 de ce tableau sont établies sur une échelle allant de 1 (le plus important) à 6 (le moins important).
1 désigne le but le plus important du formulaire.
6 désigne le but le moins important du formulaire.

Le pourcentage des moniteurs qui ont alloué le code 1 au but "feedback", a augmenté après l'utilisation du formulaire. La standardisation de la note a été perçue par contre comme un but moins important par la moitié des moniteurs qui lui allouent le code 5.

En dernier lieu, la présentation matérielle du formulaire (8 pouces et demie sur 14 pouces) semblait influencer négativement bon nombre de cliniciens et d'étudiants.

Les modifications apportées

Tenant compte de ces résultats, le comité a alors modifié le formulaire original. Le plus difficile fut la reformulation des objectifs. Il a semblé opportun de procéder à des regroupements d'objectifs. Les termes ou les phrases que les cliniciens et les étudiants ont jugés obscurs ont été revus et longuement discutés. Un lexique s'est alors avéré nécessaire et il fut élaboré par la suite.

Les commentaires sur la qualité de l'atteinte des objectifs confirmaient une faiblesse déjà constatée par les membres du comité. Après discussion, on établit une cotation de l'atteinte des objectifs en cinq degrés: non atteint, en progrès, atteint, maintenu, non applicable dans le milieu ou dans le contexte du stage. Restait à débattre la difficile question des objectifs obligatoires et complémentaires et celle de la note accordée à l'étudiant. Depuis la création du comité, les membres s'interrogeaient sur l'idée de combiner dans un même formulaire l'évaluation formative et l'évaluation normative. Au cours de l'expérimentation, les résultats révélaient que l'aspect normatif du formulaire risquait d'entraver l'aspect formatif. On demanda donc l'autorisation d'évaluer un stage clinique en terme de «réussite» ou «échec», ce que ne permettaient pas les règlements pédagogiques jusqu'à maintenant. Cette autorisation fut accordée rendant ainsi possible l'élimination de l'aspect normatif de l'évaluation et favorisant l'aspect formatif du formulaire. De plus, la présentation du formulaire a été modifiée; son format est de 8 pouces sur 11 pouces.

Les implications futures

Au cours du trimestre d'automne 1977, une expérimentation du formulaire modifié a été entreprise avec une population similaire à celle de la première expérimentation. Les réactions et les commentaires des cliniciens et des étudiants seront recueillies à l'aide d'un questionnaire plus précis que le premier.

Après l'analyse des résultats, on procédera à une étude comparative des données des deux expérimentations. Le comité, par la suite, apportera les modifications nécessaires.

Eventuellement, une étude approfondie des objectifs cotés non applicables permettrait, selon le cas, soit d'éliminer des objectifs non réalistes, soit d'aider le centre à améliorer son enseignement clinique.

Conclusion

Un travail d'une telle envergure ne peut être entrepris sans la collaboration intensive des personnes en cause et l'appui soutenu d'un agent du Service pédagogique. Fréquemment, on omet, dans ce genre de travail, le groupe le plus directement concerné par le sujet, soit les étudiants qui sont la raison d'être de la formation clinique.

D'un formulaire où tout était subjectif ou presque, le comité en a élaboré un nouveau visant l'objectivité. Afin d'améliorer la qualité de l'évaluation de la formation clinique un formulaire plus adéquat ne suffit pas à lui seul, il faudrait peut-être aussi penser à offrir aux moniteurs de stage la possibilité d'augmenter leurs qualités d'évaluateur. Différentes solutions peuvent être envisagées telles que: les services d'un professeur-coordonnateur qui pourrait assister les coordonnateurs de formation clinique des différents centres pour améliorer leur programme d'enseignement et des cours, s'adressant aux moniteurs de stage, sur la supervision d'étudiants.

Même après un travail assidu de deux ans, le comité demeure conscient que la solution apportée à l'évaluation de la formation clinique nécessite des changements adaptés à l'évolution constante de la formation de l'étudiant.

Notes et références bibliographiques

Ce texte a déjà été publié dans *Pédagogiques*, Montréal, Service pédagogique, Université de Montréal, Vol. 3 No 2, Avril 1978, pp. 31-33.

1.Blaymaker, J.E. Crocker, L.M. Muthard, J.E., *Field work Performance Report Manual*. American Occupational Therapy Association, 1974.

ÉVALUATION DE LA FORMATION CLINIQUE EN ERGOTHERAPIE: 2e ET 3e année*

L'étudiant devra être capable de:	Non atteint	En progrès	Atteint	Maintenu	Non Applicable	
1. COMPILATION DES DONNEES:					M	C
1.1 Recueillir les informations nécessaires avant l'évaluation du patient						
1.1.1. Choisir les rapports pertinents à lire et en donner les raisons						
1.1.2 Dresser le tableau clinique du patient (symptômes, phase de la maladie, pronostic,...)						
1.1.3 Identifier les traitements (passés et actuels) reçus par le patient						
1.1.4 Identifier les problèmes du patient justifiant une demande d'évaluation en ergothérapie						
1.2 Avant ou après l'évaluation, recueillir de l'information supplémentaire et en justifier l'importance						
1.2.1 Communiquer avec les sources de référence, les parents ou toute autre source de références (intra ou extra-muros)						
1.2.2 Comparer le fonctionnement du patient dans le milieu clinique et dans son milieu de vie						
2. ÉVALUATION DU PATIENT:						
2.1 Planifier l'évaluation (standardisée ou non standardisée) du patient						
2.1.1 Nommer les divers schèmes de référence et les diverses méthodes d'évaluation tout en tenant compte des informations recueillies (cf. 1.1, 1.2						
2.1.2 Analyser les avantages et désavantages de ces schèmes de référence et de ces diverses méthodes d'évaluation						
2.1.3 Sélectionner un schème de référence et une ou des méthodes d'évaluation et adapter ces dernières, s'il y a lieu						
2.1.4 Prévoir le temps et la séquence appropriés pour administrer la ou les évaluations(s)						

Ceci n'est qu'une partie du formulaire d'évaluation.

	Non atteint	En progrès	Atteint	Maintenu	Non Applicable	
					M	C
2.1.5 Préparer physiquement et psychologiquement le patient pour l'évaluation						
2.1.6 Choisir un espace adéquat tant au plan physique que psychologique						
2.1.7 Vérifier et préparer l'équipement et le matériel avant l'arrivée du patient						
2.2 Administrer l'évaluation (standardisée ou non standardisée)[1]						
2.2.1 Participer à une évaluation administrée par l'ergothérapeute						
2.2.2 Sous supervision, évaluer le patient						
2.2.3 Inscrire les résultats						
2.2.4 Interpréter les résultats (forces et faiblesses du patient)						
2.2.5 Confirmer l'exactitude de l'interprétation des résultats auprès du patient ou de sa famille						
3. PLANIFICATION DU TRAITEMENT:						
3.1 Élaborer un plan de traitement						
3.1.1 À partir des résultats de l'évaluation, réviser le choix du schème de référence						
3.1.2 Déterminer les buts à long terme du traitement ergothérapique et donner les raisons du choix de ces buts						
3.1.3 Établir les objectifs à court et moyen terme du traitement et en donner les raisons						
3.1.4 Tenir compte de l'opinion du patient et de sa famille, de la pathologie et de la personnalité du patient quant aux buts visés dans le traitement						

1. Ceci est la première évaluation de l'étudiant mais pas nécessairement la première évaluation de ce patient.